Developing cyanobacteria as a microbial factory for the production of novel compounds

Yuen Tin Lui

Department of Structural and Molecular Biology

UCL

A thesis submitted for the degree of Doctor of Philosophy

March 2017

Declaration

‘ I, Yuen Tin Lui confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Abstract

Fossil fuels are a finite resource and sustainable alternative methods of producing fuel molecules and the numerous chemical products derived from fossil fuels need to be found. Genetically engineered cyanobacteria represent one promising 'green' platform for the production of biofuels and high-value chemicals. As photosynthetic prokaryotes, cyanobacteria are able to fix atmospheric carbon dioxide using sunlight and channel that carbon into organic products. Manipulation of these biochemical pathways should allow the hyper-accumulation of certain products or the synthesis of novel compounds. Compared to other phototrophic platforms (plants and eukaryotic algae), cyanobacteria are much easier to grow and easier to manipulate genetically. Nevertheless, the current genetic tools to engineer cyanobacteria are not as advanced as those developed for *E. coli* and *S. cerevisiae*, and there is a need for new technologies in this area. The work presented in this thesis therefore aimed to develop novel tools for cyanobacteria and demonstrate their use as a platform for the production of new compounds.

The research focussed on the model organism *Synechocystis* sp. PCC 6803 – the availability of a sequenced and annotated genome and the ability to integrate DNA via natural transformation at a specific locus by homologous recombination making it ideal for genetic engineering. A new method was developed in *Synechocystis* for creating transformants that no longer contain an antibiotic resistance marker, and improvements to transcription and translation of transgenes investigated. The expression of GFP and limonene synthase were tested under various promoters, and several native and non-native integenic regions, containing ribosomal binding sites, were functionally tested in a two-gene operon containing GFP as a reporter gene. The production of limonene and (*S*)-styrene oxide, a chiral compound used as a building block for pharmaceuticals, was attempted in *Synechocystis* using the newly developed tools.

Acknowledgments

I would like to acknowledge and extend my gratitude to those who have helped me over the past 4 years.

First of all I would like to thank the BBRSC for funding my PhD. I would like to thank my supervisor, Saul Purton for giving me the opportunity to work in his group. I am extremely grateful for all the help and advice he has given. Thanks also go to John Ward and Frank Baganz for their useful advice during my thesis committee meetings.

I thank all those in the Purton group; it has been a pleasure to work with such a wonderful group of people. They have made this experience much more enjoyable and I really appreciate all their assistance during my time in the lab. In particular, I would like to thank Lamya Al-Haj and Tommasso Barbi for their guidance during the early stages of my PhD. I would also like to thank Rosie Young and Priscilla Rajakumar for their help and advice. Thanks also go to Thushi Sivagnanam for organising and running the lab. Thanks also go to Laura Stoffels, Marco Larrera Alvarez, Xenia Spencer-Milnes, Juliana Da Costa Ramos, Umaima Al Hoquanni and Max Blanshard for their company and encouragement.

I would like to extend my appreciation to Rosie, Andrew Hartley, Janet Waterhouse and Lee Murray for proof reading parts of my thesis and thank Chris Grant for his time and help using the GC-MS.

I would like to thank my mom, dad, sister and brother for all their support.

Finally, I would like to thank Lee for all his support and encouragement. Thanks for being awesome!

Abbreviations

5-FC	5-fluorocytosine
5'UTR	5' untranslated region
ADP	Adenosine diphosphate
Anabaena 7120	*Anabaena* sp. strain PCC 7120
asRNA	antisense RNA
ATP	Adenosine triphosphate
BG-11	Blue-Green 11
bp	Base pair
CCM	Carbon Concentrating Mechanism
Chl *a*	Chlorophyll *a*
Chl *b*	Chlorophyll *b*
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRi	CRISPR interference
crRNA	*cis*-repressive mRNA
cyt b_6f	Cytochrome b_6f
DMAPP	Dimethylallyl pyrophosphate
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
Fd	Ferredoxin
G.O.I.	Gene of interest
GC-MS	Gas chromatography-mass spectrometry
GFP	Green fluorescence protein
GGPP	Geranylgeranylpyrophosphate
GPP	Geranyl diphosphate
GPPS	geranyl diphosphate synthase
GT	Glucose tolerant
HA	Human influenza hemagglutinin
HPLC	High performance liquid chromatography
IG	Intergenic region
IncQ	Incompatibility group Q
IPP	Isopentenyl pyrophosphate

IR	Infrared
kDa	Kilodalton
Km^R	kanamycin resistance gene
LB	Lysogeny broth
LS	Limonene synthase
MEP	Methylerythritol 4-phosphate
mRNA	Messenger RNA
Ms	*Mentha spicata*
MVA	Mevalonate
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NEB	New England Biolabs
OD	Optical density
PAL	Phenylalanine ammonia lyase
PC	Plastocyanin
PCR	Polymerase chain reaction
PHB	Poly-β-hydroxybutyrates
PQ	Plastoquinone
PSI	Photosystem I
PSII	Photosystem II
RBS	Ribosome binding site
RNA	Ribonucleic acid
RNAP	RNA polymerase
RuBisCO	Ribulose 1,5-bisphosphate carboxylase/oxygenase
S. cerevisiae	*Saccharomyces cerevisiae*
SD	Shine-Dalgarno
SD	Standard deviation
SDS	Solenesyl diphosphate synthase
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SMO	Styrene monoxygenase
SO	Styrene oxide
Synechococcus 7942	*Synechococcus elongatus* PCC 7942
Synechococcus 7002	*Synechococcus* sp. PCC 7002
Synechocystis	*Synechocystis* sp. PCC 6803
taRNA	*trans*-activating RNA

tCa	*trans*-cinnamate
TFBS	Transcription factor binding site
WT	Wild-type
zsk	Zeocin, spectinomycin and kanamycin
σ	Sigma

Table of contents

List of figures

List of tables

Chapter 1 Introduction

1.1 Cyanobacteria

1.1.1 Overview

Cyanobacteria (formerly known as blue-green algae) are photoautotrophic bacteria capable of oxygenic photosynthesis. It is thought that cyanobacteria were the first organisms capable of oxygenic photosynthesis and played a major role in altering the earth's reducing atmosphere into the oxygenic atmosphere that is present today (Rasmussen et al., 2008). The oldest fossils date cyanobacteria at 3.5 billion years old and were discovered in western Australia (Schopf and Packer, 1987). The chloroplasts from plants and algae are also widely thought to have descended from cyanobacteria via endosymbiosis (McFadden, 2001). Cyanobacteria still play an important role today; marine cyanobacteria have a huge part to play in the carbon cycle and strains capable of nitrogen fixation play an important role in the nitrogen cycle (Capone, 2001).

The morphology of cyanobacteria varies, with certain species being unicellular or filamentous and some capable of forming heterocysts and of fixing nitrogen (Figure 1.1). Cyanobacteria do not have flagella, but some strains are able to glide. Cyanobacteria represent a diverse phylum of bacteria, and this is reflected by the wide distribution of environments they inhabit. They have been found almost everywhere, from oceans to hot springs and deserts to the Antarctic (Ruffing, 2011). The cyanobacteria that are capable of surviving in some of these harsh environments have adapted in order to do so. Cell division occurs by binary fission in cyanobacteria (Ikeuchi and Tabata, 2001).

Cyanobacteria are classified as Gram-negative bacteria, with a cell wall made up of a plasma membrane, peptidoglycan layer and outer membrane (Liberton et al., 2006). Also, most cyanobacteria contain thylakoid membranes, an internal membrane system that follows the periphery of the cell (Figure 1.2) (Vermaas, 2001). Both the photosynthetic and respiratory electron transport chains are located in the thylakoid membrane (Vermaas, 2001).

Figure removed

Figure 1.1 Microscopy images of various cyanobacteria strains.

(A) Synechocystis sp. PCC 6803, (B) *Gleotrichia echinulata* colony, (C) *Microcystis* sp., (D) *Planktothrix suspense*, (E) *Lyngbya* sp., (F) *Synechococcus elongatus* PCC 7942. Photographs from (Ruffing, 2011).

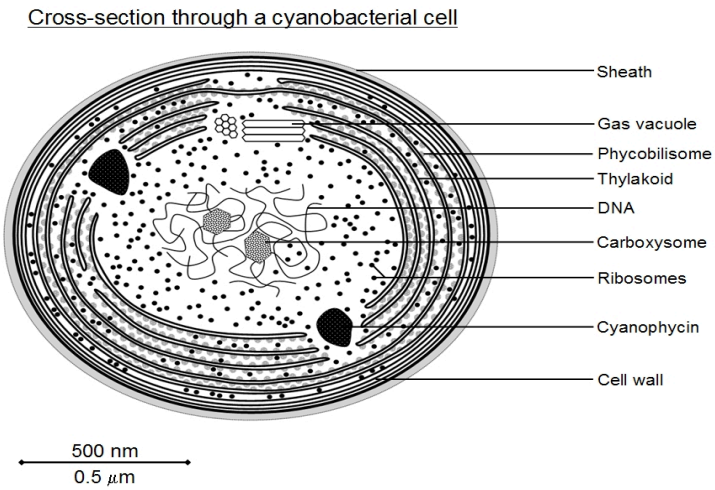


Figure 1.2 Diagram of a cyanobacteria cell.

Showing the internal membrane system, thylakoid membrane present within cyanobacteria. Diagram taken from http://cronodon.com/BioTech/Cyanobacteria.html.

In cyanobacteria the photosynthetic electron transport chain is very similar to plants, containing both photosystem II (PSII) and photosystem I (PSI) (Figure 1.3). Unlike plants, most cyanobacteria contain phycobilisomes - an antennae protein complex attached to the thylakoid membrane made up of phycobiliproteins. Allophycocyanin and phycocyanin and depending on the species, phycoerythrin, phycoerythrocyanin are the phycobiliproteins that make up the phycobilisome (Bryant, 1982; Grossman et al., 1993). The phycobilisome absorbs light at different wavelengths from the chlorophylls and carotenoids in order to increase the spectral range of light harvested (Figure 1.4)(Lea-Smith et al., 2014).

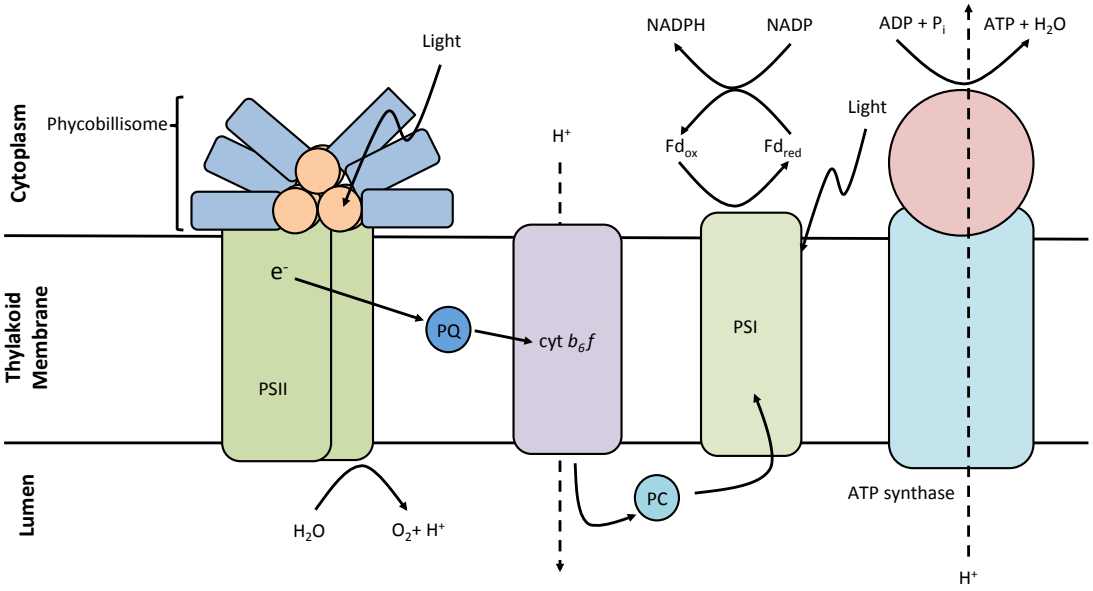


Figure 1.3 Schematic representation of the components involved in the light reaction of photosynthesis.

PSII uses light energy to split water and reduce the plastoquinone (PQ) pool; the electrons are then transferred to the cytochrome b_6f complex (cyt b_6f). Electrons are then accepted by a soluble electron carrier, plastocyanin (PC), which reduces the oxidised PS I reaction centre chlorophyll. The oxidised chlorophyll in PSI is formed when light causes the transfer of an electron to ferredoxin (Fd), which results in the production of NADPH. NADPH is used along with ATP, formed from the proton gradient created from the electron transport chains, for CO_2 fixation.

Figure removed

Figure 1.4 Absorbance spectra of the main pigments found in cyanobacteria

Phycoerythrin and phycocyanin cover different wavelengths to chlorophyll *b* (Chl *b*) and chlorophyll a, (Chl *a*). Image taken from http://www.wur.nl/en/Show/Natural-food-colorants-from-microalgae-1.htm.

The production of NADPH and ATP from the light reaction drives the fixation of atmospheric CO_2, through the Calvin cycle and the pentose phosphate pathway. The enzyme responsible for CO_2 fixation, ribulose 1,5-bisphosphate carboxylase/oxygenase

(RuBisCO) is present in the carboxysomes, which are bacterial compartments found in cyanobacteria. Carboxysomes play a central role in the carbon concentrating mechanism (CCM), which involves the uptake of inorganic carbon into the cell and the accumulation of HCO_3^-. HCO_3^- is able to diffuse into the carboxysome, and is converted back to CO_2 by carbonic anhydrase so that RuBisCO can use it as a substrate. The carboxysome was thought to have evolved in order to minimise RuBisCO reacting with O_2 by creating a barrier, a wasteful reaction that competes with CO_2 as the substrate, (Kerfeld and Melnicki, 2016).

1.1.2 *Synechocystis* sp. PCC 6803

Synechocystis sp. PCC 6803 (*Synechocystis* from hereafter) is a species of cyanobacteria found ubiquitously in freshwater that is unicellular and coccoid in shape but is unable to fix nitrogen (Figure 1.5) (Ikeuchi and Tabata, 2001). The strain was widely used to study photosynthesis as it can naturally uptake DNA from the media and is capable of growing in heterotrophic conditions using glucose (Grigorieva and Shestakov, 1982; Williams, 1988). In 1996, *Synechocystis* was the first photoautotroph and third prokaryote to have its genome fully sequenced (Kaneko et al., 1996). The genome of *Synechocystis* is organised into the main chromosome (3.57 Mbp) and seven smaller plasmids and has an average GC content of 47.7% (Kaneko et al., 1996). In *Synechocystis* it has been reported that each cell has 8-12 copies of the genome, however one publication suggested that the copy number may be much higher, with 142 copies of the genome present in the exponential phase (Griese et al., 2011; Labarre et al., 1989). The most recent publication on ploidy in *Synechocystis* suggests that at an OD: 0.1, there are around 20 copies, and the copy number varies with the growth stage and external conditions (Zerulla et al., 2016). A database was created, CyanoBase[1], to easily access the genome of *Synechocystis* (Nakamura et al., 1998). The database has been annotated and so far 3317 genes have been identified on the main chromosome. An extension of CyanoBase was created called CyanoMutants[2], a repository for genotypic and phenotypic details of mutants created in *Synechocystis* (Nakamura et al., 1999).

Synechocystis sp. PCC 6803 was originally isolated from a freshwater pond in Oakland, California in the 1970's, and is also known as the Berkeley strain. However there are now several distinct strains that originate from the Berkeley strain, that are all often referred to as *Synechocystis* sp. PCC 6803. The Berkeley strain branched into a motile strain (PCC 6803) and a non-motile strain (ATCC 27184), which are both sensitive to glucose. A

[1] http://genome.microbedb.jp/cyanobase/sequence
[2] http://genome.microbedb.jp/cyanobase/mutants/

glucose tolerant (GT) strain was created from ATCC 27184 (Ikeuchi and Tabata, 2001; Williams, 1988). The Kazusa strain used to obtain the genome sequence in 1996 is derived from the GT strain (Kaneko et al., 1996). The PCC strains were separated based on their phototactic movement into PCC-P (positive phototaxis) and PCC-N (negative phototaxis). A spontaneous mutation in the GT strain after being maintained under photoautotrophic conditions led to a strain (WL) that was no longer able to grow under mixotrophic conditions (Hihara and Ikeuchi, 1997). There is also a GT and motile 'Moscow' strain of *Synechocystis* called PCC-M (Trautmann et al., 2012).

Synechocystis can grow in the absence of a fixed carbon source (photoautotrophic growth) but in order to do so, PSII and PSI must be functional. If glucose or another fixed-carbon source is present *Synechocystis* is also capable of growing with either a functional PSI or PSII (photoheterotrophic growth) or with both PSI and PSII functional (photomixotrophic growth) (Vermaas, 1996). *Synechocystis* can also grow in light-activated heterotrophic conditions, which means that it can grow on glucose in the dark as long as it is pulsed with white light for five minutes each day (Anderson and McIntosh, 1991). *Synechocystis* has a doubling rate of 8-12 hours in photoautotrophic conditions (Heidorn et al., 2011; Yu et al., 2013).

Figure removed

Figure 1.5 Images of *Synechocystis* 6803

(A) Photograph of *Synechocystis* culture in liquid and solid medium (taken by Michelle Liberton, Washington University from http://www.enerzine.com/6/9788+produire-du-biocarburant-par-les-microorganismes-i+.html). (B) Light micrograph image *Synechocystis* 6803 from the Protist Information Server[3] (C) Electron microscope transmission image of *Synechocystis* (adapted from (Nickelsen et al., 2011)).

[3] http://protist.i.hosei.ac.jp/pdb/Images/Prokaryotes/Chroococcaceae/Synechocystis/index.html

1.1.3 Other model organisms of cyanobacteria

Currently there are 39 genomes of different cyanobacteria species on CyanoBase that are sequenced and annotated, however only a few species are widely used along with *Synechocystis* as model organisms representing cyanobacteria (Nakamura et al., 1998). *Synechococcus elongatus* PCC 7942 (*Synechococcus* 7942) is a species of cyanobacteria that has been widely used to study circadian rhythms (Holtman et al., 2005). It is a unicellular, freshwater strain like *Synechocystis*, and was the first strain to be naturally transformed (Shestakov and Khyen, 1970). The genome of *Synechococcus* has also been sequenced. The filamentous, heterocyst-forming cyanobacteria, *Anabaena* sp. strain PCC 7120 (*Anabaena* 7120) has also been used a model of cyanobacteria to study nitrogen fixation (Rice et al., 1982). *Anabaena* 7120 is a freshwater strain that can be transformed via conjugation with *E. coli* and electroporation and was fully sequenced in 2001 (Kaneko et al., 2001; Moser et al., 1995; Wolk et al., 1984).

The thermophilic, unicellular, rod-shaped cyanobacterial strain, *Thermosynechococcus elongatus* BP-1, has been used for protein structural studies as the proteins are more stable at a higher temperature and thus more likely to give rise to stable crystals suitable for X-ray crystallography (Iwai et al., 2004; Onai et al., 2003). *Thermosynechococcus elongatus* BP-1 has been sequenced and can be transformed naturally, and by electroporation and conjugation (Mühlenhoff and Chauvat, 1996; Nakamura et al., 2002; Onai et al., 2003)

1.1.4 Applications of cyanobacteria

The exploitation of engineered cyanobacteria originally focused on their ability to grow in certain environmental niches. For example, cyanobacteria were thought to be excellent vectors to deliver mosquitocidal toxins, as they are found in the breeding grounds for mosquitoes and are eaten by the larvae (Ruffing, 2011). *Anabaena* 7120 was found to be a suitable host for the expression of mosquitocidal toxins for mass production; it is capable of expressing mosquitocidal toxin genes from *Bacillus thuringiensis subsp. israelensis* and was shown to still be larvicidal 4 years after the construction of the strains (Lluisma et al., 2001).

Cyanobacteria has also been engineered for the potential use as biosensors to measure the major limiting nutrients such as iron, phosphorus and nitrate that lead to toxic algae blooms, in order to predict their occurrence (Ruffing, 2011). The biosensors would often involve the expression of the *luxAB* reporter genes under a nutrient responsive promoter such as the *isiAB* iron sensitive promoter that was tested in *Synechococcus* 7942 to detect

the levels of iron in Lake Huron, Canada (Porta et al., 2003). Cyanobacteria has also been engineered to detect herbicides and heavy metals using the expression of the luciferase, *luc,* gene from firefly as the reporter (Shao et al., 2002). One issue with using cyanobacteria as a biosensor is the promoter and the concentration it detects the target compound.

More recently, there has been a lot of interest in engineering cyanobacteria for production of small molecule organics such as ethanol and *iso*-butanol that could be used as biofuels. This is further discussed in the next section.

1.2 Production of chemicals using modified cyanobacteria

1.2.1 Overview

The world is heavily dependent on fossil fuels with these accounting for 81.2% of all energy production (International Energy Agency, 2016a). It is recognised that fossil fuels such as coal, oil and gas are finite resources. However, fossil fuel consumption continues to increase. With the global energy demand predicted to increase by 37% by 2040 from 2014 levels (International Energy Agency, 2014), there is a greater urgency to find alternative sources of energy to cope with demand in the future.

Oil is the most used fuel and accounts for 32.9% of global energy consumption (BP, 2016; International Energy Agency, 2016b). Although solar, wind and wave can be used to provide a renewable source of electrical energy and can help to reduce greenhouse gas emissions, renewable liquid fuels such as biofuels are considered the only replacement for transport fuels (i.e. petrol, diesel and jet fuel) unless there is a complete overhaul of the transport infrastructure (Machado and Atsumi, 2012; Quintana et al., 2011). Of all the oil consumed in 2013, 63.8% was used for transport and 16.2% was for non-energy uses (International Energy Agency, 2015). Most chemical feedstocks used for the production of everyday products such as plastics, polystyrene and pharmaceuticals are also derived from oil. Therefore alternative and sustainable methods for the production of these chemicals are also needed (Erickson et al., 2012). Engineered microorganisms could provide an alternative 'bio-manufacturing' technology for synthesis of these chemicals. To-date, the majority of work has focused on the model organisms *Escherichia coli* (*E. coli*) and *Saccharomyces cerevisiae* (*S. cerevisiae*), as their genetics are well-characterised, a wide range of genetic engineering tools are available and they are facultative anaerobes with fast growth rates (Lee et al., 2008). However, one of the biggest issues with using *S. cerevisiae* and *E. coli* is that they both require a carbon source, usually sugar, which comes from plants. These plants then compete directly for land and resources with crops that are required for food. Sugars obtained from breaking down non-edible crop materials, i.e. lignocellulose, could be utilised instead however a viable method for the production of low-value chemicals such as bioethanol and biomethane has not yet been found (Machado and Atsumi, 2012).

Photosynthetic microorganisms such as algae and cyanobacteria have also generated interest as a potential platform for the production of organic chemicals, due to their ability to fix atmospheric CO_2 using sunlight. Other potential advantages to using photosynthetic microorganisms is the removal of atmospheric CO_2, one of the main greenhouse gases

responsible for global warming, and the ability of photosynthetic microorganisms to grow on non-arable land (Machado and Atsumi, 2012; Ruffing, 2011). They also grow much faster than plants and per hectare are expected to produce a higher yield (Nozzi et al., 2013; Quintana et al., 2011).

Algae are eukaryotic organisms and some species, when stressed, are capable of producing large amounts of neutral lipids that can be converted into biodiesel (Mubarak et al., 2015). However the genetic engineering of algae for the production of novel compounds, is much more difficult in comparison to cyanobacteria. As prokaryotes, cyanobacteria have smaller genomes than eukaryotic algae, and their genomes lack introns and extensive intergenic regions. For some cyanobacterial species there are established protocols for the transfer of DNA, together with characterised plasmid vectors, selection markers and gene expression systems (Elhai, 1994; Koksharova and Wolk, 2002; Vioque, 2007). Other advantages to using cyanobacterial species are their ability to grow in extreme environments and their simple nutrient requirements (Ruffing, 2011). Furthermore, N_2-fixing species do not require costly nitrogen fertilisers that crops need, which are a known source of nitrous oxide, which is a greenhouse gas that has a damaging effect on the ozone layer (Machado and Atsumi, 2012).

1.2.2 Chemical production in cyanobacteria

As shown in Table 1.1, cyanobacteria have been genetically modified to produce a number of chemicals at lab-scale (Machado and Atsumi, 2012; Ruffing, 2011). The majority of the work has been achieved using *Synechococcus* 7942 and *Synechocystis*, due to the availability of an annotated reference genome, and well-established molecular techniques for genetic manipulation (Rosgaard et al., 2012). Both species are naturally transformable and will readily import exogenous DNA into the cell and efficiently integrate the DNA into the genome via homologous recombination making genetic engineering quick, easy and predictable (see section 1.3 for further details) (Koksharova and Wolk, 2002). However, these strains were not selected for their ability to produce novel compounds and other strains with superior traits including the ability to fix nitrogen, produce hydrogen, grow quickly and accumulate biomass, may be more suitable for the production of novel compounds. *Leptolyngbya* sp. strain BL0802 was identified as a strain with traits that would make it ideal for commercial production, such as growth at a wide range of temperatures and tolerance to high levels of light and salt. The strain was also capable of accumulating lipids and monounsaturated fatty acids as a biodiesel feedstock (Taton et al., 2012).

Table 1.1 Biofuels and high-value products produced in engineered strains of cyanobacteria.

Chemical	Use	Cyanobacteria strain	Reference
Ethanol	Biofuels	*Synechocystis* 6803 *Synechococcus* 7942	(Dexter and Fu, 2009) (Deng and Coleman, 1999)
1-Butanol	Biofuels	*Synechococcus* 7942	(Lan and Liao, 2011)
Alkanes/ Alkenes	Biofuels	*Synechocystis* 6803	(Tan et al., 2011)
Ethylene	Platform chemicals	*Synechococcus* 7942 *Synechocystis* 6803	(Sakai et al., 1997; Takahama et al., 2003) (Ungerer et al., 2012)
Isobutanol	Biofuel and platform chemical	*Synechococcus* 7942 *Synechocystis* 6803	(Atsumi et al., 2009) (Varman et al., 2013)
Isobutyraldehyde	Biofuel and platform chemical	*Synechococcus* 7942	(Atsumi et al., 2009)
Isoprene	Biofuel and platform chemical	*Synechocystis* 6803	(Lindberg et al., 2010)
2,3-butanediol	Biofuel and platform chemical	*Synechococcus* 7942 *Synechocystis* 6803	(Oliver et al., 2013) (Savakis et al., 2013)
Fatty acids	Biofuel	*Synechocystis* 6803 *Synechococcus* 7942	(Liu et al., 2011a) (Kaiser et al., 2013)
Fatty Alcohols	Biofuel	*Synechocystis* 6803	(Tan et al., 2011)
Squalene	Cosmetics, vaccines, potential feedstock chemicals	*Synechocystis* 6803	(Englund et al., 2014)
Lactic acid	Biofuel and platform chemical	*Synechocystis* 6803 *Synechococcus* 7942	(Angermayr et al., 2012) (Niederholtmeyer et al., 2010)
2-methyl-1-butanol	Biofuel	*Synechococcus* 7942	(Shen and Liao, 2012)
Sucrose	Feedstock for bioethanol fermentation	*Synechocystis* 6803 *Synechococcus* 7942	(Du et al., 2013) (Ducat et al., 2012)
Glycerol	Platform chemical	*Synechocystis* 6803	(Savakis et al., 2015)
***β*-carophyllene**	Fragrance and cosmetics	*Synechocystis* 6803	(Reinsvold et al., 2011)
***β*-phellanderene**	Fragrance	*Synechocystis* 6803	(Bentley et al., 2013a)
Mannitol	Food and pharmaceutical industry	*Synechococcus* sp. PCC 7002	(Jacobsen and Frigaard, 2014)
Isopropanol	Platform chemical	*Synechococcus* 7942	(Kusakabe et al., 2013)
1,2-Propanediol	Antifreeze, cosmetics, plastics	*Synechococcus* 7942	(Li and Liao, 2013)
1,3-Propanediol	Platform chemical	*Synechococcus* 7942	(Hirokawa et al., 2016)
Acetone	Platform chemical	*Synechocystis* 6803	(Zhou et al., 2012)
Cellulose	Feedstock for bioethanol fermentation	*Synechococcus leopoliensis* strain UTCC 100 *Synechococcus* sp. PCC 7002	(Nobles and Brown, 2008) (Zhao et al., 2015)

Chemical	Use	Cyanobacteria strain	Reference
Glycogen	Feedstock for bioethanol fermentation	*Synechococcus* sp. PCC 7002	(Aikawa et al., 2014)
Farnesene	Biofuels and fragrances/ cosmetics/ lubricants	*Anabaena* 7120	(Halfmann et al., 2014a)
3-hydroxybutyrate (PHB)	Plastic	*Synechococcus* 7942	(Suzuki et al., 1996)
		Synechocystis 6803	(Tyo et al., 2009)
		Synechococcus sp. MA19	(Nishioka et al., 2001)
		Synechococcus sp. PCC 7002	(Zhang et al., 2015)

1.2.2.1 Biofuels and biofuel feedstocks

Biofuels are defined as fuels that are derived from biomass, including any organic material from life forms or its metabolic products (Quintana et al., 2011). Two commercially available biofuels are bioethanol, produced from corn starch or sugar cane, and biodiesel produced from vegetable oils and fats. However, both of these biofuels are not ideal replacements for existing fuels, as bioethanol is difficult to transport, has only 70% the energy density of petrol and is hygroscopic (Nozzi et al., 2013; Peralta-Yahya et al., 2012). Biodiesel is also difficult to transport due to problems of solidification at low temperatures, is not as energy dense as diesel and in cold weather biodiesel can also clog up fuel lines in vehicles (Peralta-Yahya et al., 2012).

The focus for 'next generation biofuels' is therefore advanced fuel molecules produced by genetically engineered microorganisms such as cyanobacteria. These biofuels would be hydrocarbons that have similar properties to petroleum-based fuels and can be used directly as 'drop-in' fuels with existing engine technologies. Such fungible biofuels include compounds such as butanol and higher alcohols, linear and cyclic alkane and alkenes, and branched isoprenoid hydrocarbons (Table 1.2).

Table 1.2 Table of the components present in current transport fuels, and the ideal properties required when choosing alternatives. Table from (Lee et al., 2008).

Fuel Type	Major Components	Important Properties	Biosynthetic alternatives
Petrol	C_4-C_{12} Hydrocarbons Linear, branched, cyclic, aromatic Anti-knock additives	Octane Number Energy Content Transportability	Ethanol, *n*-butanol and *iso*-butanol Short chain alchols Short chain alkanes
Diesel	C_9-C_{23} Hydrocarbons (Average C_{16}) Linear, branched, cyclic, aromatic Anti-freeze additives	Cetane number Low Freezing Temperature Low vapour pressure	Biodiesel (FAMEs) Fatty alcohols, alkanes Linear or cyclic isoprenoids
Jet Fuel	C_8-C_{16} Hydrocarbons Linear, branched, cyclic, aromatic Anti-freeze additives	Very low freezing temperature Net heat of combustion Density	Alkanes Biodiesel Linear or cyclic isoprenoids

One of the first attempts of biofuel production in cyanobacteria was in 1999 when ethanol was produced in *Synechococcus* 7942 after expressing genes for pyruvate decarboxylase (*pdc*) and alcohol dehydrogenase (*adh*) from *Zymomonas mobilis* (Deng and Coleman, 1999). Since then, ethanol has also been produced in *Synechocystis* by expressing the same genes under the control of the *psbAII* promoter (Dexter and Fu, 2009). A number of the biofuels produced in cyanobacteria have utilised the natural metabolic pathways already present, only requiring the additional expression of a few genes in most cases (Figure 1.6) (Ruffing, 2011). Fatty acids, fatty alcohols, alkanes and alkenes are just a few potential biofuels that been produced in cyanobacteria (Kaiser et al., 2013; Liu et al., 2011a; Tan et al., 2011)

Figure removed

Figure 1.6 Metabolic pathways present in cyanobacteria utilised for the production of biofuels and other high-value products.

The solid arrows indicate enzymes that are naturally found, whereas the dashed arrow requires the expression of a recombinant enzyme. In the dashed boxes are the products that have been or could be produced in cyanobacteria. Figure modified from (Ruffing, 2011). Abbreviation: GPP – geranyl diphosphate; IPP – isopentenyl pyrophosphate; DMAPP – dimethylallyl pyrophosphate; HMBPP - 1-hydroxy-2-methyl-2-butenyl-4-pyrophosphate; Me-CPP - 4-diphosphocytidyl-2-C-methylerythritol; CDP-MEP - 4-diphosphocytidyl-2-C-methyl-D-erythritol 2-phosphate; CDP-ME - 2-C-methyl-D-erythritol 2,4-cyclodiphosphate; MEP - 2-C-methylerythritol 4-phosphate; DXP - 1-deoxy-D-xylulose 5-phosphate ; GAP - glyceraldehyde 3-phosphate; 3-PGA - 3-phosphoglyceric acid; RuBP - ribulose-1,5-bisphosphate; F6P - fructose 6-phosphate; Glc - glucose; PPP - pentose phosphate pathway; PEP – phosphoenolpyruvic acid; PYR – Pyruvate; AA – amino acids; 2-AL – 2-acetolactate; 2,3-DHIV -2,3-dihydroxy-isovalerate; 2-KIV - 2-ketoisovalerate; Val – valine; IBA- isobutyraldehyde ; IBOH - isobutanol ; OAA - oxaloacetic acid; 2-OG – 2-oxoglutarate; acetyl-CoA – acetyl coenzyme A ; acyl-ACP – acyl-acyl carrier protein ; acyl-A – acyl-aldehyde ; FFA – free fatty acids.

1-Butanol production in *Synechocystis* 7942 was successfully achieved by expressing five genes, three of which are from *Clostridium acetobutylicum,* a bacterium capable of naturally producing 1-butanol (Lan and Liao, 2011). Butanol is thought to be a possible

replacement for gasoline, as it is more energy dense than ethanol and has low hygroscopicity (Lan and Liao, 2011). Isobutyraldehyde is a useful chemical feedstock as it can be converted into a number of products such as isobutanol, a potential replacement for petrol that is usually derived from oil. Both isobutyraldehyde and isobutanol have been successfully produced in *Synechococcus* 7942. Isobutanol has also been produced in *Synechocystis* after expressing two genes, to produce a yield of 298 mg/L under mixotrophic culture conditions (Varman et al., 2013).

The production of biofuel feedstocks in cyanobacteria has also been investigated. Cellulose, a natural product of plants, can be used as a feedstock for microorganisms for the production of biofuels once it has been broken down into smaller sugars. The issue with plant-produced cellulose is that it is associated with lignin and hemicellulose, making it difficult to break down. Cellulose is also a component of some cyanobacterial cell walls and its over-production by genetic engineering yields lignin-free cellulose from sunlight and CO_2 that could potentially be used for biofuels and other biotechnological applications (Nobles and Brown, 2008; Zhao et al., 2015). The over-production and secretion of simple sugars such as sucrose, glucose and fructose that are naturally made in cyanobacteria, has also been achieved and can be utilised as carbon source for other microbes for the production of biofuels (Du et al., 2013; Ducat et al., 2012; Niederholtmeyer et al., 2010).

1.2.2.2 High-value chemicals

With the average price of oil in 2015 at its lowest level since 2004 (BP, 2016), the production of biofuels cannot compete with the low cost of oil. There are also major questions over whether the production of biofuels in cyanobacteria can be economically feasible (Ducat et al., 2011a; Ruffing, 2011). Therefore, the current focus is on the production of high-value chemicals in cyanobacteria, whilst the technology for biofuels continues to be developed (Ducat et al., 2011a). Some of the high-value products that have been produced in cyanobacteria are discussed further below. An early attempt was made in *Synechocystis* to overexpress native genes that encode enzymes involved in carotenoid biosynthesis as there is commercial interest in the use of these natural pigments and anti-oxidants in the food, cosmetic, nutraceutical and pharmaceutical industry (Lagarde et al., 2000).

Ethylene is one of the most widely used chemicals; it is used for the production of plastics such as polyethylene and polyester, and is used to produce industrial grade ethanol (Ungerer et al., 2012). Ethylene is produced from steam cracking long chain hydrocarbons from petroleum and this results in the production of significant amounts of CO_2 waste. The

production of ethylene was achieved in *Synechococcus* 7942 by introducing the *efe* gene from *Pseudomonas syringae*, which encodes an ethylene-forming enzyme that converts 2-oxoglutarate into ethylene and carbon dioxide (Takahama et al., 2003). However, the strain producing ethylene was not genetically stable (Takahama et al., 2003). Changes were made to *efe* before the production of ethylene was achieved in *Synechocystis* (Ungerer et al., 2012).

Isoprene is a platform chemical used to produce a number of other high value chemicals. It can be polymerised to make a synthetic rubber, and there is also interest in using it as a biofuel (Ducat et al., 2011a; Wang et al., 2012). Isoprene is produced in plants from dimethylallyl pyrophosphate (DMAPP), which is synthesised via either the methylerythritol 4-phosphate (MEP) pathway or the mevalonate (MVA) pathway (Figure 1.7). Like other prokaryotes, cyanobacteria contain only the MEP pathway, and by introducing an isoprene synthase into *Synechocystis*, isoprene was successfully produced (Lindberg et al., 2010). The yield of isoprene was further improved 2.5 fold, by introducing a heterologous MVA pathway into *Synechocystis* (Bentley et al., 2013b).

The success of isoprene production has drawn interest into the production of terpenes and terpenoids in cyanobacteria. Terpenoids, often referred to as isoprenoids, are modified terpenes. Terpenes are made up of units of isoprene (C_5H_8), although isopentenyl pyrophosphate (IPP) and DMAPP are the building blocks for terpenes. IPP like DMAPP is produced from either the MEP or MVA pathway. Both terpenes and terpenoids have useful applications in medicines, fragrances, flavourings, cosmetics and as potential biofuels (Pattanaik and Lindberg, 2015). Squalene is a 30-carbon isoprenoid that is used in the cosmetic industry and has potential use as a replacement for petrol. It is naturally produced in *Synechocystis*, but is usually converted to hopene. A genetically engineered strain inactivating the gene predicted to encode squalene hopene cylase led to the accumulation of squalene, 70 times higher than levels seen in wild-type cells (Englund et al., 2014).

Poly-β-hydroxybutyrates (PHB) are thought to be promising replacements for plastics, which are currently produced from fossil fuels (Ruffing, 2011). One of the major advantages of using PHB and other polyhydroxyalkanoates as replacements for plastic is that they are biodegradable, making these materials a more environmentally friendly product (Samantaray et al., 2014). PHB is naturally produced in some strains of cyanobacteria and usually the yield is quite low. However, *Synechococcus* sp. MA19 naturally accumulates PHB up to 27% of dry cell weight (Miyake et al., 2000; Panda and Mallick, 2007). Attempts to increase the production of PHBs, by altering the culture

conditions demonstrated that growth in media deficient in nitrogen and phosphorus resulted in increase accumulation (Panda and Mallick, 2007).

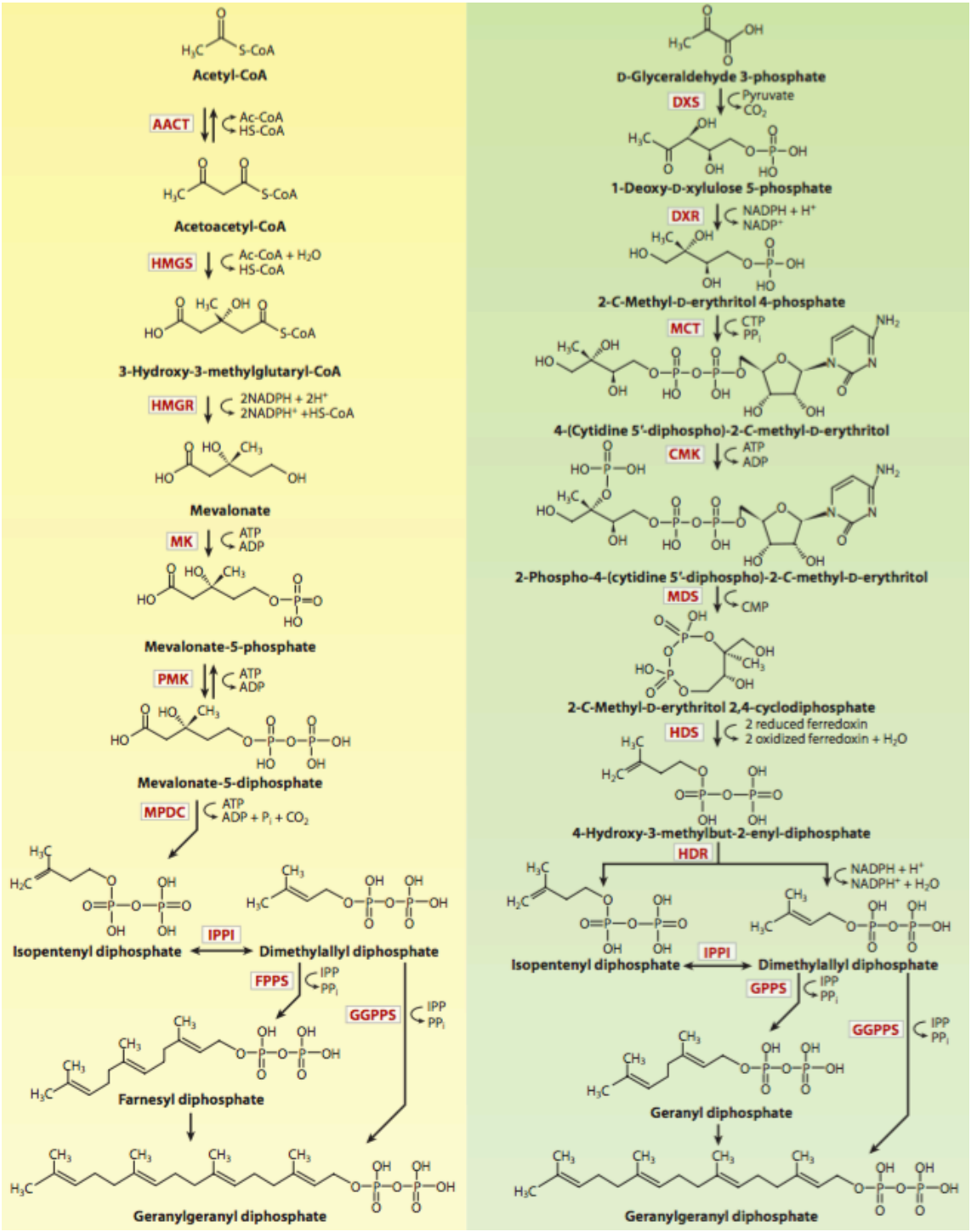


Figure 1.7 MVA and MEP pathway present in plants.

Figure from (Vranová et al., 2013).The MVA pathway (in yellow) is not present in cyanobacteria but the MEP pathway (in green) is present and can be utilised for the production of terpenes and terpenoids.

1.2.3 Challenges and opportunities

Although cyanobacteria have been successfully engineered to produce biofuels and other chemicals at the lab-scale there are still significant challenges that need to be overcome before commercialisation (Parmar et al., 2011). There are some technical challenges, regarding cultivation and harvesting, and ecological challenges to ensure the environmental benefits of using cyanobacteria are achieved. There are also economic challenges related to the cost of scaling up production (Parmar et al., 2011). This section will focus on some of the challenges faced by producing novel compounds in cyanobacteria and the potential opportunities for improvement through genetic engineering.

1.2.3.1 Photosynthesis

As photosynthesis is driving the production of biofuels and high-value products in cyanobacteria, improving the photosynthetic efficiency when growing cells in dense, large-scale monocultures has the potential to improve yield. Improving light harvesting CO_2 fixation and the CCM in cyanobacteria are all strategies that have been investigated.

When attempting to produce large-scale cultures of cyanobacteria under phototrophic growth conditions, the light-harvesting pigments can become a problem due to excessive photon capture by the cells at the surface that then block light availability to the cells below (Wang et al., 2012). The absorbed excess light energy is dissipated as heat and can result in reduced photosynthesis by photoinhibition, due to damage to the photosynthetic proteins by high light (Ruffing, 2011). Attempts have been made to reduce the antenna size (Cotton et al., 2015). However, *Synechocystis* strains that were engineered to be deficient in phycobilisomes, were not useful for improving the production of biomass unless carbon was limited (Lea-Smith et al., 2014).

RuBisCO, the enzyme responsible for fixing CO_2 is thought to be the rate-limiting step to carbon fixation (Jensen, 2000). When isobutyraldehyde production from CO_2 was successfully achieved in *Synechococcus* 7942, the production was further increased by expressing the *rbcLS* genes encoding RuBisCO, from the related strain *S. elongatus* strain PCC6301 (Atsumi et al., 2009). In this case the rate of O_2 produced from photosynthesis did not increase so the increased carbon fixation may be due to a more efficient use of the reducing power generated from photosynthesis. In *Synechocystis*, a mutant RuBisCO was introduced, that had an improved activity but less specificity for CO_2 over O_2, however there was no growth improvement and less RuBisCO was present in the cells (Durão et al., 2015).

1.2.3.2 Toxicity

One of the major problems with attempting to produce biofuels and high value products in microorganisms is product toxicity. Therefore, it is important to try and alleviate any toxic effects as this could limit production (Dunlop, 2011). Most of the work on product toxicity has focussed on biofuel toxicity, particularly towards *E. coli* and *S. cerevisiae* (Dunlop, 2011). Not all biofuels are toxic, and the antimicrobial effect of biofuels is thought to relate to how hydrophobic the molecule is, as this determines how much accumulates in the cytoplasmic membrane (Brennan et al., 2012; Dunlop, 2011). Most biofuels interfere with the cell membrane and result in increased permeability of the cell, which then affects physiological processes. Some biofuels may interfere with membrane proteins, altering the fluidity of the membrane, which then affects the stability and structure of the cell (Dunlop, 2011).

Attempts have been made to reduce the toxic effect of biofuels with some success. The heterologous expression of an efflux pump in *E. coli* improved survival in the presence of limonene and also increased levels of production (Dunlop et al., 2011). In *S. cerevisiae*, the toxic effect of limonene was alleviated using non-toxic solvents such as dibutyl phthalate, dioctyl phthalate, isopropyl myristate and farnesene acting as an organic phase over the aqueous medium (Brennan et al., 2012). This strategy could also help in the extraction and recovery of the biofuels. Overexpression of heat shock proteins and molecular chaperones that prevent protein aggregation during heat shock, could also be employed to reduce the toxic effect of biofuels (Dunlop, 2011).

1.2.3.3 Improving the molecular techniques required for metabolic engineering.

Product yields could be further improved in engineered cyanobacteria by applying synthetic biology and systems biology approaches. This is discussed further in the next section, along with the current tools that are available in cyanobacteria.

1.3 Genetic engineering of cyanobacteria

As detailed in the previous sections, a few strains of cyanobacteria have been genetically engineered for the production of biofuels and high-value products. The molecular tools available to introduce foreign DNA into cyanobacteria and engineer changes in endogenous genes, have been developed for a number of model strains as follows:

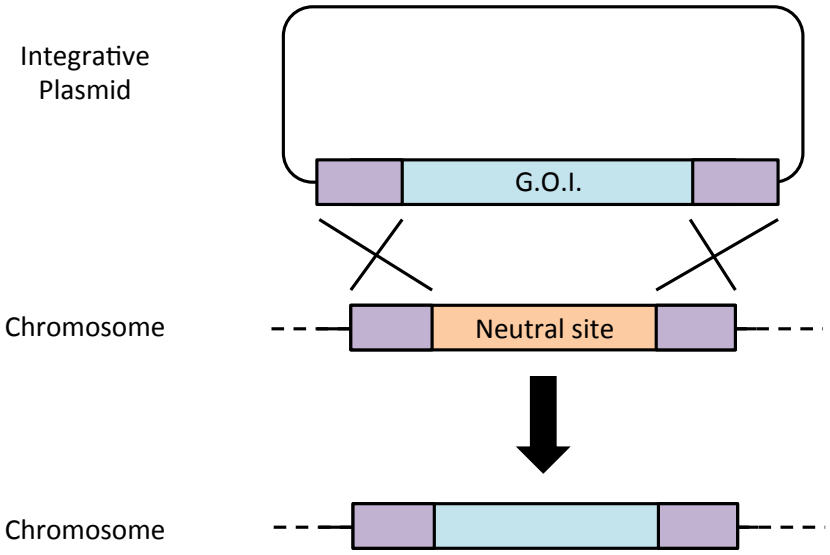
1.3.1 Transformation of cyanobacteria

There are three methods that can be used to introduce DNA into cyanobacteria; natural transformation, electroporation and conjugation.

The ability of *Synechocystis* to naturally take up DNA from the medium and efficiently integrate this DNA into the genome by homologous recombination is one of the main reasons why this species is used as a model organism for cyanobacteria (Grigorieva and Shestakov, 1982; Ikeuchi and Tabata, 2001). Unicellular cyanobacteria such as *Synechocystis*, *Synechococcus* 7942, *Synechococcus* PCC 7002 (*Synechococcus* 7002) and *Thermosynechococcus elongatus* BP-1 are just some of the strains of cyanobacteria that are naturally transformable. With natural transformation, DNA (on a plasmid or as linear DNA) is naturally taken up from the media. DNA can then integrate into the genome by double homologous recombination, which is favoured compared to single homologous recombination (Figure 1.8) (Heidorn et al., 2011). The mechanism of natural transformation is not clear, although it is thought to involve DNA in its single stranded-form in *Synechocystis* (Barten and Lill, 1995). In *Synechocystis*, the gene *slr0197* is thought to encode a DNA-binding protein and is required for DNA uptake along with the presence of pili (Yoshihara et al., 2001). After finding homologs of pili and other genes vital for natural transformation in *Synechocystis* present in *Thermosynechococcus elongatus* BP-1, a method was developed for naturally transforming this strain (Iwai et al., 2004).

In *Synechocystis*, natural transformation was demonstrated to be the most efficient method in comparison to conjugation and electroporation (Zang et al., 2007). Transformation success was improved when cells were transformed during mid log phase and when the cells were incubated with DNA for 5 hours before plating onto antibiotic free media to allow cells to recover (Zang et al., 2007). Transformation with circular DNA was also found to be more successful than using linearized DNA (Kufryk et al., 2002).

(A) Double homologous recombination



(B) Single homologous recombination

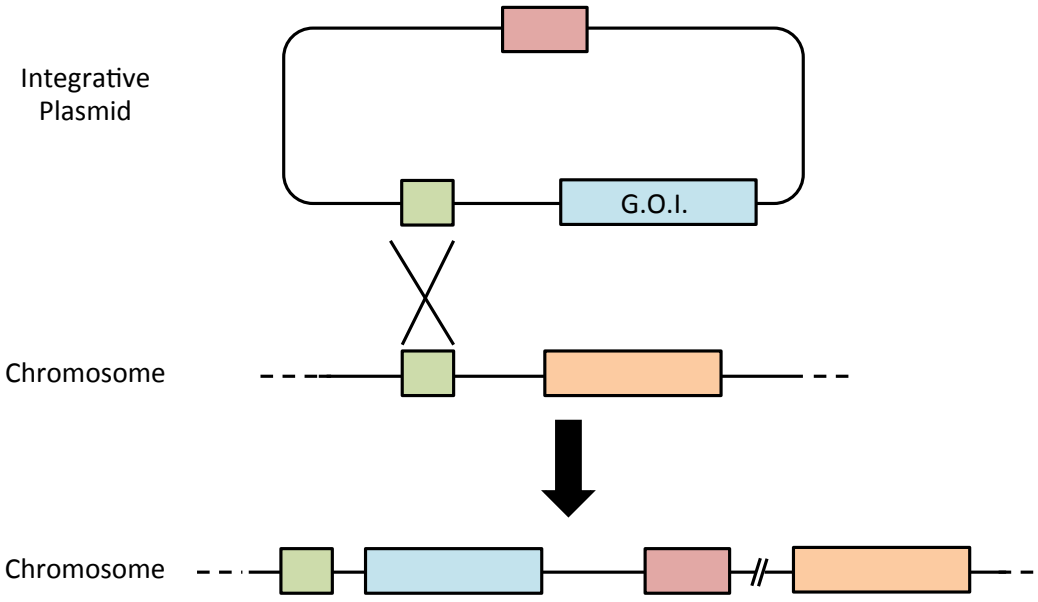


Figure 1.8 Foreign DNA can be inserted into the cyanobacteria genome by (A) double homologous recombination or (B) single homologous recombination.

In (A) the G.O.I. replaces the WT gene and in (B) the entire plasmid integrates into the genome.

Both conjugation and electroporation can be used to introduce DNA into cyanobacteria strains that are not naturally transformable. These two methods can also be used when single recombination is required or when introducing a replicative plasmid (Heidorn et al., 2011).

Conjugation with *E. coli* is often used in filamentous cyanobacteria strains such as *Anabaena* and *Nostoc* strains and can be used in the unicellular strains *Synechococcus* 7942, *Synechococcus* PCC 6301 and *Synechocystis* (Flores and Wolk, 1985; Mühlenhoff and Chauvat, 1996; Tsinoremas et al., 1994; Wolk et al., 1984). Conjugation usually uses two strains of *E. coli* (Figure 1.9). One strain contains the conjugative plasmid that carries the genes that enable the transfer of DNA. The other strain contains the cargo plasmid and the helper plasmid. The cargo plasmid contains the DNA to be introduced into the cyanobacteria and also the oriT site. The *mob* gene is also required as it encodes a nickase that recognises the oriT site and enable the transfer of the cargo plasmid into cyanobacteria, although this can be provided *in trans* on the helper plasmid. The helper plasmid also contains genes for DNA methylases that protect the foreign DNA from being degraded by the host cyanobacteria's restriction enzymes (Elhai, 1994; Elhai and Wolk, 1988; Vioque, 2007). The cargo plasmid can be a replicative plasmid or contain DNA that is capable of integrating by recombination or transposition (see Plasmid vectors for more details).

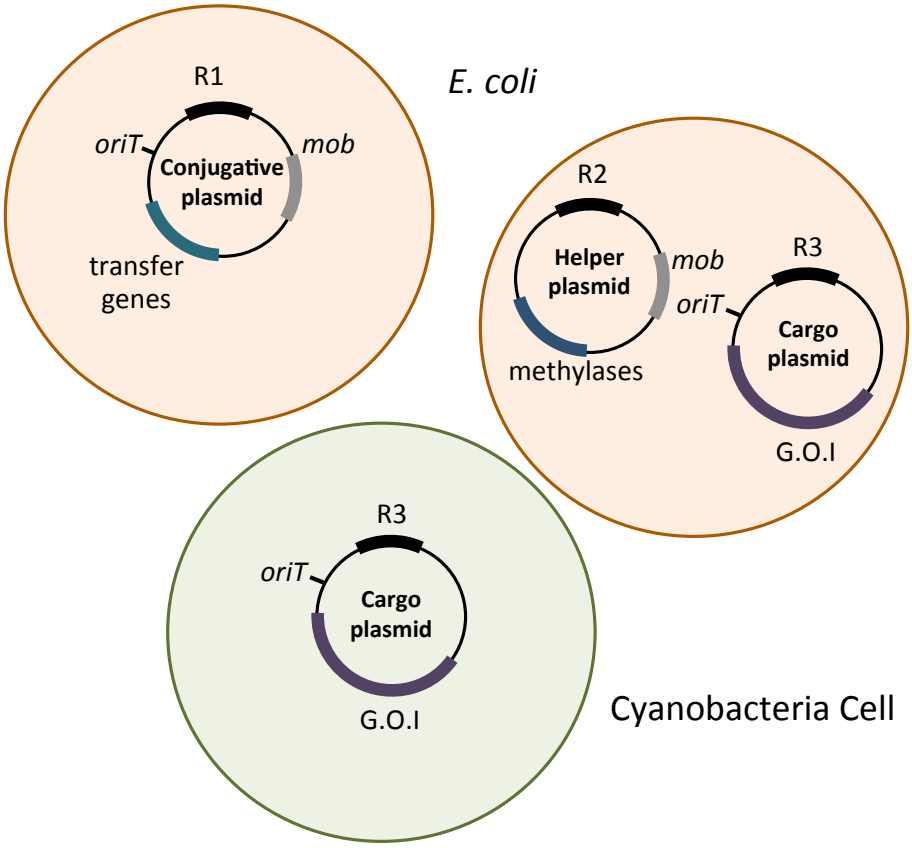


Figure 1.9. Triparental mating procedure used to introduce DNA into cyanobacterial cells by conjugation.

G.O.I., Gene of interest. Figure was adapted from (Vioque, 2007).

In *Anabaena* PCC 7120, double recombination events using conjugation do not take place however the use of the *sacB* gene, which is lethal when cells are grown in the presence of sucrose, can eventually result in the production of double recombinants (Cai and Wolk, 1990).

Electroporation can also be used to introduce a plasmid into some strains of cyanobacteria, such as *Anabaena* (Thiel and Poo, 1989). Electroporation works by sending a pulse of electricity in order to open pores in the cell membrane for DNA to enter. Electroporation is also possible for freshwater and marine *Synechococcus* strains, and can be improved with $CaCl_2$ treatment (Matsunaga and Takeyama, 1994). One of the issues with this method is that is may lead to mutations in the genome (Elhai, 1994).

1.3.1.1 Segregation

Some cyanobacteria contain multiple copies of the chromosome in each cell (Griese et al., 2011; Labarre et al., 1989). As aforementioned the copy number in *Synechocystis* varies greatly depending on the growth phase and the environment (Zerulla et al., 2016). The number of genome copies also varies in *Synechococcus* dependent on the growth phase

(Watanabe et al., 2015). Therefore when foreign DNA is inserted into a copy of the genome by homologous recombination, continued rounds of selection may be required to segregate the genome to the homoplasmic state, where all copies contain the insertion and no wild-type copies remain. Once all copies of the genome contain the gene, the selection pressure no longer needs to be maintained. If complete segregation is not possible, and copies of the insertion and the wild-type persist, this state is referred to as persistent heteroplasmy, and suggests that the wild-type version of the targeted locus is essential for cell viability (Vermaas, 1996).

1.3.2 Plasmid vectors

There are two types of plasmid vector that are used to introduce DNA into cyanobacteria: integrative and replicative vectors. Integrative vectors carry foreign DNA that can then integrate into the cyanobacterial genome by homologous recombination, whereas replicative vectors are able to replicate in the host and are passed onto the daughter cells. They are therefore able to maintain foreign DNA without integration into the main chromosome (Heidorn et al., 2011).

1.3.2.1.1 Replicative plasmids

Since molecular cloning takes place in *E. coli*, a shuttle vector that is able to replicate in both *E. coli* and the cyanobacteria host is required. There are two types of shuttle vectors that are used in cyanobacteria: those that contain a broad host replicon, and those that contain an *E. coli* cloning plasmid replicon and a cyanobacteria-specific replicon. Examples of the latter type are: i) pSCR119, which contains the replicon pDC1 from *Nostoc* sp. MAC 8009 and the pMB1 *E. coli* replicon; ii) pUC303, which contains the pUH24 replicon from *Synechocystis* 7942 and the p15A *E. coli* replicon (Kuhlemeier et al., 1983; Summers et al., 1995). Broad-host range plasmids usually contain the replicon derived from the incompatibility group Q (IncQ) plasmid RSF1010. Shuttle vectors pFC1, pSL1211 and pPMQAK1 are a few examples of replicative plasmids and all contain the RSF1010 replicon (Wang et al., 2012). The pPMQAK1 has been demonstrated to be functional in *Synechocystis*, *Anabaena* 7120 and *Nostoc* ATCC29133 (Huang et al., 2010). To maintain the presence of the replicative plasmids, antibiotic selection is required. The copy-number of the plasmid needs to be considered when selecting a suitable plasmid. Plasmids with the RSF1010 replicon have been demonstrated as having 10 copies in *E. coli* and 10 – 30 copies in *Synechocystis* (Marraccini et al., 1993; Ng et al., 2000).

When expressing genes on a replicative plasmid, antibiotic selection is often required to maintain the presence of the plasmid. However, maintaining the antibiotic selection for a

large culture can become costly. In *Synechococcus* 7002, a method was established, where antibiotics would not need to be used, to maintain the replicating plasmid by complementing a lethal *recA* null mutant with an *E. coli recA* gene present on the plasmid (Akiyama et al., 2011).

1.3.2.1.2 Integrative vectors

Integrative vectors are non-replicating plasmids and are therefore lost from the cyanobacterial population during cell division, so the DNA needs to integrate into the genome. Any plasmid can be used to transfer DNA into the host strain. Integrative vectors insert DNA into the genome at a specific location by homologous recombination. DNA can integrate by single or double homologous recombination depending on the cyanobacteria strain (Vioque, 2007). The host-homologous DNA elements flanking the introduced DNA determine the recombination site. The minimum length of these elements differs for different strains of cyanobacteria, although it is generally thought that increasing the flanking region increases the efficiency of recombination (Zang et al., 2007).

When DNA is introduced into the genome the target site chosen is often a neutral site that does not negatively affect the strain. This could be an intergenic site or could involve the disruption (or replacement) of a non-essential gene. The *psbAII* gene is often used as a neutral site in *Synechocystis* as the deletion of this gene, which encode a photosystem II D1, does not affect growth as *psbAIII*, which encodes the same protein, is overexpressed to compensate for the deletion (Mohamed et al., 1993). The *psbAI* gene which encodes the same protein has also been used as a neutral site (Varman et al., 2013). The *slr0168* gene encodes a hypothetical protein in *Synechocystis* that has also been used as a neutral site (Bentley et al., 2013b). One issue with using intergenic regions is that they may actually be transcribed or have antisense function (Mitschke et al., 2011). More neutral sites in *Synechocystis* are being discovered to enable the expression of heterologous pathways at multiple sites (Ng et al., 2015).

Transposons have also been used in some strains of cyanobacteria e.g. *Nostoc* ATCC29133 and *Synechococcus* 7942 in order to introduce DNA into the genome without the use of recombination (Elhai, 1994; Heidorn et al., 2011).

1.3.3 Selectable markers

Selection for successful transformants is usually made by introducing an antibiotic resistance gene along with the G.O.I. and choosing the transformants that are capable of growing in the presence of the antibiotic. Chloramphenicol, neomycin, streptomycin, kanamycin, spectinomycin, ampicillin, bleomycin and gentamycin resistance cassettes can

all be used in different cyanobacteria for selection (Ruffing, 2011). Antibiotics that are light sensitive such as tetracycline and rifampicin are less useful in cyanobacteria as they are usually incubated under light (Koksharova and Wolk, 2002). A method of selection using herbicide resistance towards bialaphos and phosphinothricin has also been developed in *Synechocystis* and *Synechococcus* 7942 (Chungjatupornchai, 1990; Takahashi et al., 1999).

1.3.4 Reporters

Numerous reporter proteins have been used in cyanobacteria as a method to investigate gene expression and protein localisation.

In *Synechocystis* the *luxAB* genes, from *Vibrio harveyi* (*V. harveyi*), encoding luciferase have been expressed along with *luxCDE*, also from *V. harveyi*, which produces the substrate luciferase aldehyde that is required to produce light (Peca et al., 2008). The use of *luxAB* has also been utilised in cyanobacteria without the expression of the *luxCDE*, where the substrate is provided in the media (Gillor et al., 2003; Muramatsu and Hihara, 2006). As mentioned earlier, the luciferase, *luc*, gene from the firefly *Photinus pyralis* has also been utilised in *Synechocystis* (Shao et al., 2002).

Although the presence of photosynthetic pigments in cyanobacteria can interfere with fluorescence signals, some fluorescent proteins have been used as reporter genes. GFP and derivatives of GFP e.g. Cerulean, GFPmut3B and EYFP have all been used successfully (Huang et al., 2010), although one of the issues with using GFP is that photobleaching can be a problem (Koksharova and Wolk, 2002).

The *lacZ* gene, from *E. coli*, encodes β-galactosidase, which has also been used in cyanobacteria as a reporter. The β-galactosidase activity assay in cyanobacteria is performed with an alternative fluorescent substrates instead of Xgal, as the blue product is not suitable due to the colour of cyanobacteria (Scanlan et al., 1990; Thiel et al., 1995). The *cat* gene encodes chloramphenicol acetyl transferase, which result in chloramphenicol resistance. The enzyme transfers the acetyl group from acetyl-coA onto chloramphenicol, preventing it from binding and inhibiting the ribosome. The expression of *cat* has been used as a reporter in various strains of cyanobacteria by measuring resistance to chloramphenicol (Bauer and Haselkorn, 1995; Marraccini et al., 1993).

1.3.5 Biological parts in cyanobacteria

There is growing interest in applying synthetic biology principles to cyanobacteria, in order to use them as 'green' platforms for the production of biofuels, foods, feed and other

platform chemicals (Lindblad et al., 2012; Wang et al., 2012). The main aim of synthetic biology is to try and genetically modify organisms in a predictive way by designing and making new biological parts, devices and systems (and also redesigning existing biological systems), for the production of useful products (Huang et al., 2010). Examples of biological parts are promoters, ribosome binding sites (RBSs), terminators and coding sequences. By having various well-characterised parts, the organism can be easily engineered and parts can be swapped in and out if necessary to optimise the yield (Heidorn et al., 2011). So far, the majority of work on synthetic biology has focussed on *E. coli* however, some of the parts developed in *E. coli* do not function in the same way in cyanobacteria (Heidorn et al., 2011). This section describes some of the biological parts that have been developed in some of the model strains so far. However, to further utilise the part-based approach to construct and control new metabolic pathways, more biological parts need to be characterised for use in cyanobacteria.

1.3.5.1 Promoters

Transcription is the first step of gene expression and in prokaryotes it is a step involving a RNA polymerase (RNAP) holoenzyme, a complex that consist of a sigma (σ) subunit (σ factor) and the core RNAP enzyme made up of the subunits α_2 (RpoA x 2), β (RpoB), β' (RpoC) and ω (RpoZ). σ factors play an important role in controlling transcription, as different σ factors will bind to the core enzyme depending on the environmental conditions and different σ factors recognise different classes of promoter. In cyanobacteria the RNAP holoenzyme and the set of σ factors are different to that seen in *E. coli*; the β' subunit is split into two different subunits RpoC1 (ϒ)) and RpoC2 (β'), which has a DNA binding domain at the C-terminus. In other bacteria the σ factors belong to the σ^{70} family or the σ^{54} family, however, in cyanobacteria σ factors in the σ^{54} family are absent (Imamura and Asayama, 2009).

In cyanobacteria, both native and foreign promoters have been used to express foreign genes (Table 1.3). The native promoters used to express foreign genes constitutively in cyanobacteria are often promoters of genes that are important for photosynthesis e.g. P_{psbAII} (photosystem II D1 protein), P_{rbcL} (Rubisco large subunit), P_{cpc} (c-phycocyanin beta subunit) and P_{psaD} (photosystem I subunit II) (Wang et al., 2012). The *psbAII* promoter is seen as a constitutive promoter in *Synechocystis*, as *Synechocystis* is usually grown in the presence of light. Expression is seen in low light however, the promoter can also serve as an inducible promoter as expression can be increased when grown in the presence of high light (500 μmol photons/s/m^2) (Lindberg et al., 2010). A super strong constitutive promoter, P_{cpc560}, was recently discovered; made up of two *cpcB* promoters and 14

predicted transcription factor binding sites (TFBSs). Using this promoter, expression of heterologous genes were reported at levels achieved in *E. coli* (Zhou et al., 2014).

Table 1.3 Table of promoter used to express heterologous genes in cyanobacteria

Promoter	Source	Constitutive/ Inducible [4]	Examples – Host gene		Reference
P_{psbAII}	*Synechocystis*	Constitutive	*ispS*	*Synechocystis*	(Lindberg et al., 2010)
P_{psbAI}	*Synechococcus 7942*	Constitutive	*efe*	*Synechococcus 7942*	(Sakai et al., 1997)
$P_{psbAIII}$	*Synechocystis*	Constitutive	*lhcB*	*Synechocystis*	(He et al., 1999)
P_{rbc}	*Synechococcus* 6301	Constitutive	luciferase	*Synechococcus* 6301	(Takeshima et al., 1994)
	Synechococcus 7942	Constitutive	*pdc*	*Synechococcus* 7942	(Deng and Coleman, 1999)
	Synechocystis	Constitutive	*far*	*Synechocystis*	(Tan et al., 2011)
P_{rpnB}	*Synechocystis*	Constitutive	GFPmut3B	*Synechocystis*	(Huang et al., 2010)
P_{nrsB}	*Synechocystis*	Inducible	holin	*Synechocystis*	(Liu and Curtiss, 2009)
P_{cpc}	*Synechocystis*	Constitutive	*accB and accC*	*Synechocystis*	(Liu et al., 2011b)1)
P_{cpc560}	*Synechocystis*	Constitutive	*ispS*	*Synechocystis*	(Zhou et al., 2014)
P_{psaA}	*Synechocystis*	Constitutive	*luxAB*	*Synechocystis*	(Muramatsu and Hihara, 2006)a)
P_{psaD}	*Synechocystis*	Constitutive	*luxAB*	*Synechocystis*	(Muramatsu and Hihara, 2007)
P_{cmp}	*Synechocystis*	Inducible	*fol*	*Synechocystis*	(Liu et al., 2011b)1)
P_{sbt}	*Synechocystis*	Inducible	*gpl*	*Synechocystis*	(Liu et al., 2011b)1)
P_{lac}	*E. coli*	Inducible	*hydA*	*Synechococcus* 7942	(Ducat et al., 2011b)1)
P_{trc}/ P_{tac}	*E. coli*	Inducible	*petE*	*Synechococcus* 7942	(Geerts et al., 1995)
P_{coaT}	*Synechocystis*	Constitutive (Repressed in presence of Fe^{3+})	*efe*	*Synechocystis*	(Guerrero et al., 2012)
P_{isiAB}	*Synechocystis*	Inducible	*gfp*	*Synechocystis*	(Kunert et al., 2000)

Inducible promoters are desirable when producing a toxic compound in order to reduce the stress on the cell during growth. Many of the native inducible promoters that have been utilised are responsive to metal ions. One of the more responsive inducible promoters in *Synechocystis* is $P_{nrsB,}$ which is induced by Ni^{2+} or Co^{2+} (García-Domínguez et

[4] Grown under light conditions

al., 2000). Additionally, P_{coaT} is a native promoter induced by Co^{2+} or Zn^{2+} found in *Synechocystis* (Peca et al., 2008). Expression under P_{coaT} is not strong, but in the absence of the inducible metal, complete repression was seen (Guerrero et al., 2012). Finally, the P_{isiAB} promoter is expressed in the absence of Fe^{3+} in *Synechocystis*. The addition of Fe^{3+} then leads to repression of expression, although there is some leakage of expression (Kunert et al., 2000).

Foreign promoters, usually from *E. coli*, have also been used in cyanobacteria in order to express foreign genes. *E. coli* promoters in cyanobacteria do not function in the same way, as there are differences between the RNA holoenzyme and sigma factors in *E. coli* compared to *Synechocystis* (Heidorn et al., 2011). P_{tet} and P_{lac} are commonly used in *E. coli* but function weakly in cyanobacteria (Huang et al., 2010). Modifications have been made to certain *E. coli* promoters, e.g. P_{trc} and P_{lac}, that have helped improve their function in *Synechocystis* (Camsund et al., 2014; Huang et al., 2010).

1.3.5.2 Ribosome binding sites

Translation in bacteria is initiated when the ribosome binds to mRNA at the RBS. Within the RBS is the core Shine-Dalgarno (SD) sequence (5'-GGAGG-3'). The base pairing between the RBS with the anti-SD sequence (found on the 3' end of the 16S rRNA) and the distance between the SD sequence and the start codon affects the functioning of the RBS (Heidorn et al., 2011). The effectiveness of the RBS is also affected by secondary structures that might arise from the surrounding nucleotide sequences (Salis et al., 2009).

In *Synechocystis* the 3' end of the 16S rRNA is …AUCACCUCCUUU-3', the optimal complementary sequence is therefore AAAGGAGGUGAU (SD sequence underlined) and the ideal spacing between the central A in the SD sequence and the first base in the start codon is nine bases (Heidorn et al., 2011). The RBS sequence UAGUGGAGGU based on the optimal complementary sequence (with optimal spacing between the central A and the start codon) was compared against three BioBrick RBS and gave two-fold higher expression than the RBS sequence AUUAAAGAGGAGAAA (a strong RBS in *E. coli*) (Heidorn et al., 2011).

1.3.5.3 Codon optimisation

Codon-optimisation is often performed when introducing a gene from one organism into another to try and improve the efficiency of translation, e.g. by removing rare codons (Quax et al., 2015). In *Synechocystis*, a comparison was made to see the effect of codon-optimisation. The expression of the gene encoding isoprene synthase, *ispS*, from *Pueraria*

Montana was approximately 10 times greater after codon-optimising the gene for *Synechocystis* (Lindberg et al., 2010).

1.3.5.4 Degradation

Proteins have the potential to accumulate in the cell even after the promoter controlling the expression has been switched off, thereby having no real control over the circuit. In *E. coli*, degradation tags have been used, as proteases recognise the proteins that have been tagged. There has been some research on the use of degradation tags in cyanobacteria. A few non-native degradation tags were expressed and were functional in cyanobacteria (Huang et al., 2010). In *Synechocystis*, a range of cyanobacteria and *E. coli* degradation tags were developed and tested that provide a range of turnover levels (Landry et al., 2013).

1.3.5.5 Marker-less transformants

Transformant lines that do not carry an antibiotic resistance gene are desirable as this eliminates the risk of antibiotic resistance genes being transferred to microorganisms in the environment when growing cultures in open systems. A marker-less strategy also ensures that when performing multiple transformations on a strain, the number of antibiotic resistance cassettes that are available does not limit this. In some cyanobacterial strains the expression of the conditionally toxic *sacB* gene has been utilised to produce marker-less transformants (Cai and Wolk, 1990; Lagarde et al., 2000). In the first transformation step, the *sacB* gene alongside an antibiotic resistance cassette is introduced into the genome and selection is made based on antibiotic resistance. In the second step, both the *sacB* gene and the antibiotic selection cassette are replaced, and selection is made based on the loss of *sacB* in the toxic conditions (Ried and Collmer, 1987).

1.4 Summary and thesis aims

Fossil fuel usage continues to increase despite limited reserves and the effect of global warming. Renewable alternatives are required to produce the transport fuels and chemicals that are normally derived from oil. Cyanobacteria as photosynthetic bacteria have the potential to be a microbial factory for the production of novel compounds. The molecular tools currently available have allowed model strains of cyanobacteria to be engineered to produce novel chemicals at the lab-scale, but these tools are rather basic and lag behind the metabolic engineering technologies and synthetic biology tools of *E. coli*.

With this in mind, the main aims of the following studies were:

- Attempt the production of limonene, an essential oil produced in plants that is used in fragrances and also has potential use as a jet fuel replacement, in *Synechocystis* (Chapter 3).
- Improve the molecular tools available to cyanobacteria in order to express multiple-genes involved a metabolic pathway for the production of novel chemicals as an operon (Chapter 4), and create strains that are marker-less (Chapter 5).
- Utilise the molecular tools developed to validate their function and produce another novel compound in *Synechocystis* (Chapter 6).

Chapter 2 Materials and methods

2.1 Bacterial Strains, media and culture conditions

2.1.1 Synechocystis sp. PCC 6803

2.1.1.1 Strains used

The wild-type (WT) *Synechocystis* sp. PCC 6803 (large cells and glucose tolerant) was obtained from Prof. Conrad Mullineaux, Queen Mary University of London. The strain is referred to as WT *Synechocystis* hereafter.

Transgenic strains of *Synechocystis* sp. PCC 6803 created by previous members of the Purton laboratory were also used in the study and are listed in Table 2.1. Transgenic strains created during this study are listed in the Appendix.

Table 2.1 Transgenic strains used in this study

Strain name (in this study)	Description	Reference
6803.AII	The *psbAII* gene is replaced by no gene-of-interest under the control of the P_{psbAII} and a kanamycin resistance cassette.	(Al-Haj, 2014)
6803.nrsB	The *psbAII* promoter and gene is replaced by no gene-of-interest under the control of the P_{nrsB} and a kanamycin resistance cassette.	(Al-Haj, 2014)
6803.4op	The *psbAII* and gene is replaced by *ble, aadA, aphA-6* and *gfp* under the control of the P_{psbAII}. Resistant to zeocin, spectinomycin and kanamycin.	(Mackrow, 2011)
6803.cpc	The *psbAII* promoter and gene is replaced by no gene-of-interest under the control of the P_{cpc560} and a kanamycin resistance cassette.	(Lau, unpublished)
6803.AII.GFP	The *psbAII* gene is replaced by *gfp* under the control of the P_{psbAII} and a kanamycin resistance cassette.	(Lau, unpublished)
6803.cpc.GFP	The *psbAII* promoter and gene is replaced by *gfp* under the control of P_{cpc560} and a kanamycin resistance cassette.	(Lau, unpublished)

2.1.1.2 Media

Synechocystis was grown in BG-11 (Blue Green 11) medium. The stock solutions used for the preparation of BG-11 are presented in Table 2.2. To prepare 1L of medium: 10 ml of 100x BG-11 stock, 1 ml trace elements and 1 ml iron stock are combined with 976 ml water and autoclaved. Once the solution has cooled, 1 ml Phosphate stock, 1 ml Na_2CO_3 stock, 10 ml $NaHCO_3$ stock and 10 ml TES Buffer are added.

Table 2.2 Stock solutions for the preparation of BG-11 medium

100 x BG-11	**g/L**
$NaNO_3$	149.6
$MgSO_4.7H_2O$	7.49
$CaCl_2$	3.60
Citric acid	0.60
Na_2EDTA	1.12 ml 0.25M solution, pH 8.0
Trace Elements	**g/100ml**
H_3BO_3	0.286
$MnCl_2.4H_2O$	0.181
$ZnSO_4.7H_2O$	0.022
$Na_2MoO_4.2H_2O$	0.039
$CuSO_4.5H_2O$	0.008
$Co(NO_3)_2.6H_2O$	0.005
Iron Stock	**g/100ml**
Ferric ammonium citrate	1.11
Phosphate stock	**g/100ml**
K_2HPO_4	3.05
Na_2CO_3 stock	**g/100ml**
Na_2CO_3	2.0
TES Buffer	**g/100ml**
TES	22.9
$NaHCO_3$ stock	**g/100ml**
$NaHCO_3$	8.401g

For the preparation of solid medium, 2x concentrated BG-11 (including 3 g/L of sodium thiosulphate) was mixed with the same volume of molten 3% Difco-agar (w/v) to produce 1x BG-11 and 1.5% agar. Antibiotics (2.2) were added to the media just before pouring.

Initially BG-11 agar supplemented with 5% (w/v) sucrose was autoclaved with the 2x concentrated BG-11 solution that also contained $NaHCO_3$. BG-11 agar supplemented with 5% (w/v) sucrose was then prepared without the addition of $NaHCO_3$, and with a stock of 50% (w/v) sucrose solution that was filter sterilised using a 0.22 μm filter.

To prepare BG-11 agar supplemented with 5-fluorocytosine (5-FC), a 10 mg/ml stock of 5-FC in 2x BG-11 was prepared before adding to the rest of the 2x BG-11 medium.

2.1.1.3 Culture conditions

2.1.1.3.1 Growth of Synechocystis

Synechocystis strains were grown on BG-11 agar plates incubated at 30°C in an illuminated incubator with the light intensity between 40–80 μmol/m^2/s, unless stated otherwise.

Liquid cultures of *Synechocystis* were grown in an illuminated incubator with a light intensity of 20–200 μmol/m^2/s, at 120 rpm shaking and at 25°C, unless otherwise stated. Flasks were inoculated with a loop of cells from agar plates. Larger volumes of culture were sometimes inoculated from a smaller pre-culture, in a 1 in 10 dilution.

2.1.1.3.2 Quantification of *Synechocystis*

To quantify the growth of *Synechocystis*, 1ml of liquid culture was used to record the optical density (OD) at 730 nm using the Unicam UV/Vis Spectrometer (Thermo Electron Corporation, USA).

2.1.1.3.3 Growth of *Synechocystis* in automated photobioreactor, Algem®

In the Algem® Photobioreactor (Algenuity, UK), *Synechocystis* was grown in 200 μmol/m^2/s white light, at 30°C with 120 rpm shaking. The OD at 750 nm was recorded every 30 minutes. Cultures were prepared in 500 ml BG-11 medium, and inoculated with a starter culture to give a starting OD_{750} ~0.1. Cultures were performed in duplicate, unless otherwise stated and the average OD_{750} was plotted over time.

To calculate the growth rate: the ln (OD_{750}) between the exponential phase, OD_{750}: 0.5 and 1.5, was plotted on a graph. The gradient of the line of best fit is the growth rate and is expressed as h^{-1}.

2.1.1.3.4 Short-term storage of *Synechocystis*

Synechocystis strains were grown on BG-11 agar plates, containing antibiotic (2.2) if required and then subsequently stored in dim light conditions of 5-10 μmol/m^2/s at 20°C. Strains were transferred to a fresh plate every 3-4 weeks.

2.1.1.3.5 Long-term storage of *Synechocystis*

For the long-term storage of *Synechocystis* strains, a 50 ml liquid culture was grown to an OD_{730} between 1 and 3. Cells were harvested by centrifugation at 4500 g for 6 minutes at room temperature. The supernatant was poured off and the cell pellet resuspended in the residual medium. The volume of the cell suspension was then measured and mixed with

0.24 volumes of 80% glycerol. The cells were stored at -80 °C. The cells were recovered by inoculating BG-11 medium with a thawed stock.

2.1.2 Escherichia coli

2.1.2.1 Strain

Escherichia coli DH5α (*fhu*A2, Δ(*argF-lacZ*)U169, *pho*A, *gln*V44, Φ80, Δ(*lacZ*)M15, *gyr*A96, *recA*1, *relA*1, *endA*1, *thi*-1, *hsdR*17) was the strain used. Referred to as *E. coli* hereafter.

2.1.2.2 Media

E. coli was grown in LB (Lysogeny Broth) medium prepared with 10 g/L bacto-tryptone, 5 g/L yeast extract and 10 g/L NaCl. For the preparation of solid 1.5% Difco agar was added (15 g/L). LB medium was autoclaved before use.

Strains of *E. coli* carrying a zeocin resistance gene were grown in low-salt LB medium and 5 g/L NaCl was used instead.

2.1.2.3 Culture conditions

2.1.2.3.1 Growth of *E. coli*

E. coli cultures were grown on LB agar plates at 37°C, or in LB liquid medium under constant shaking at 37°C. Antibiotics (2.2) were added if necessary.

2.1.2.3.2 Long-term storage of *E. coli*

For long-term storage, *E. coli* cells were stored as 25% glycerol stock. Equal volumes of bacterial culture were mixed with 50% autoclaved glycerol. The mixture was then briefly vortexed and stored at –80 °C.

2.2 Antibiotics and other selective agents

A list of the antibiotics (Table 2.3) and selective agents (Table 2.4) used in this study are presented below. Unless otherwise stating the working concentration in these tables was used.

Table 2.3 List of antibiotics used in this study.

Antibiotic	Stock concentration (mg/ml)	Working concentration for *E. coli* (µg/ml)	Working concentration for *Synechocystis* (µg/ml)
Ampicillin	75 in H_2O	75	N/A
Kanamycin	50 in H_2O	50	200
Chloramphenicol	34 in 100% EtOH	34	34
Zeocin™ (Invivogen)	100 in H_2O	25	7.5
Spectinomycin	50 in H_2O	50	50

Table 2.4 List of selective agents used in this study

Selective agent	Stock concentration	Working concentration for *Synechocystis*
5- FC	10mg/ml in 2x BG-11 medium	0.5 mg/ml
Sucrose	N/A	5% (w/v)

2.3 Plasmids

Plasmids created by others used in this study are listed in Table 2.5. Plasmids created in this study can be found in the Appendix. Plasmids used for integration in *Synechocystis* are indicated with *.

Table 2.5 Plasmids used in this study

Plasmid name	Source	Usage
4op*	(Mackrow, 2011)	Expression vector for *Synechocystis*, containing *ble*, *aadA*, *aphA-6* and *gfpuv* under the control of the *psbAII* promoter for insertion at the *psbAII* locus.
4op.Rev*	(Mackrow, 2011)	Expression vector, containing *ble*, *aadA*, *aphA-6* and *gfp* under the control of the *psbAII* promoter in the reverse order of 4op for insertion at the *psbAII* locus.
pARG1.3	Saul Purton's Collection	Source of *cat* cassette.
pJET1.2 blunt	ThermoFisher Scientific™	Blunt end cloning vector; base for creating expression vectors used in this study.
pLAH.AII*	(Al-Haj, 2014)	Expression vector for *Synechocystis*, containing NdeI and BamHI restriction sites to insert a gene-of-interest (G.O.I.) under the control of the *psbAII* promoter and a kanamycin resistance cassette for insertion at the *psbAII* locus. See Appendix for a diagram.

Plasmid name	Source	Usage
pLAH.nrsB*	(Al-Haj, 2014)	Expression vector for *Synechocystis*, containing NdeI and BamHI restriction sites to insert a G.O.I. under the control of the *nrsB* promoter and a kanamycin resistance cassette for insertion at the *psbAII* locus. See Appendix for a diagram.
pLAH.cpc*	(Lau, unpublished)	Expression vector for *Synechocystis*, containing NdeI and BamHI restriction sites to insert a G.O.I. under the control of the *cpc560* promoter and a kanamycin resistance cassette for insertion at the *psbAII* locus. See Appendix for a diagram.
pRY127d	(Young and Purton, 2014)	Source of synthetically designed *codA* gene, codon optimised for *Chlamydomonas reinhardtii*.
pUC4K	(Taylor and Rose, 1988)	Source of a kanamycin resistance cassette (Km^R).
pUM24Cm	(Ried and Collmer, 1987)	Source of *sacBR –nptI* cassette.
pZΔES	Saul Purton's Collection	Source of a *ble* cassette.
pASapI-SplB	(Larrea Alvarez, unpublished)	Source of synthetically designed *splB* gene, codon optimised for *Chlamydomonas reinhardtii.*

2.4 Molecular genetic methods

2.4.1 PCR

PCR reactions were performed on a programmed thermocycler. The TC-3000X or TC-3000G thermocyclers were used (both from Techne), the latter one was also used for gradient PCR reactions. PCR was performed using Phusion® high-fidelity DNA polymerase from New England Biolabs (NEB); the components and conditions used for PCR reactions are in Table 2.6. The annealing temperature varied for the primers used.

Table 2.6 Component and cycling conditions for a 25 µl reaction mix using Phusion ® high-fidelity DNA polymerase.

Components		Volume (for a 25 µl reaction)	Final Concentration
5X Phusion Buffer		5	1x
Deoxynucleotide Solution Mix (10 mM)		0.5	200 µM
Upstream Primer (100 µM)		0.25	1 µM
Downstream Primer (100 µM)		0.25	1 µM
Template DNA		Variable	
Phusion DNA Polymerase		0.25	0.02 units/ µl
Nuclease Free Water		Add to a final volume of 25 µl	
Cycling conditions for PCR Reaction			
Stage		**Temperature (°C)**	**Time**
Initial Denaturation		98	1 minute
25 cycles	Denaturation	98	10 seconds
	Annealing	45-72	20 seconds
	Extension	72	15-30 seconds per kb
Final Extension		72	5 minutes
Storage		10	Hold

2.4.2 Primers

All primers used were synthesised by MWG Eurofins, Germany. For a list of the primers used in this study see the Appendix.

2.4.3 Agarose gel electrophoresis

DNA fragments were separated on a 1% (w/v) agarose gel made with 1 x TAE Buffer (40 mM Tris, 1mM sodium EDTA, 17.5 mM acetic acid) and 0.2 µg/ml ethidium bromide. Samples were mixed with 6x DNA Loading Dye (Thermo Scientific™) before loading into a gel that was submerged in 1 x TAE buffer in an electrophoresis tank. GeneRuler™ DNA Ladder Mix (Thermo Scientific™) was loaded alongside the samples in order to estimate the size of the DNA fragments. Electrophoresis was performed at 90 V for 30-90 minutes,

depending on the size of the expected fragments. DNA fragments were visualised using a UV transilluminator (UVP Gel Documentation System).

2.4.4 Plasmid isolation

Plasmid isolation was performed using the GeneJET Plasmid Miniprep Kit (Thermo Scientific™) according to the manufacturer's instructions.

2.4.5 DNA purification

DNA purification of PCR products and DNA fragments were performed using the GeneJET PCR Purification Kit (Thermo Scientific™) for quantities of a plasmid up to 20 µg according to the manufacturer's instructions.

2.4.6 Gel extraction

To isolate DNA from an agarose gel, the DNA of interest was excised from the gel illuminated with UV light. The DNA was purified from the gel using the GeneJET Gel Extraction Kit (Thermo Scientific™) following the manufacturer's instructions.

2.4.7 Restriction enzyme digests

Restriction enzymes from NEB or Thermo Scientific™ were used to perform restriction enzyme digests according to the manufacturer's instructions.

2.4.8 Dephosphorylation

To avoid the religation of digested vectors, dephosphorylation was sometimes used to remove the 5' phosphate. Dephosphorylation was performed using Antarctic Phosphatase (NEB). To 30 µl of digested DNA, 3 µl of 10x Antarctic Phosphatase Buffer was added along with 1 µl of Antarctic Phosphatase. After incubating the reaction mixture for 30 minutes at 37°C the reaction was heat inactivated at 65°C for 5 minutes.

2.4.9 Ligation

Ligation was performed using T4 DNA ligase (NEB) following the manufacturer's protocol. The ratio of vector to insert used for ligation varied from 1:3 to 1:5, with 50-100 ng of vector used.

2.4.10 DNA sequencing

DNA sequences of plasmids and PCR products were performed using the Scientific Support Services of the Wolfson Institute for Biomedical Research, University College

London or using Source Biosciences, Cambridge. The results were aligned using ClustalW2 or MacVector 12.6.0 software.

2.4.11 DNA quantification

To quantify the concentration of DNA fragments and plasmids, 1 µl was used for analysis on a Nanodrop 2000c Spectrophotometer (Thermo Scientific™).

2.4.12 Cloning in *E. coli*

2.4.12.1 Preparation of competent *E. coli* cells

DH5α was restreaked from a frozen glycerol stock onto a LB plate and incubated at 37°C overnight. A single colony was picked and used to inoculate 10 ml of LB and grown overnight at 37°C under continuous shaking. From this culture, 1 ml was added to 100 ml LB and grown under continuous shaking at 37°C for 2-3 hours. After incubation, the culture was cooled on ice for 15 minutes and was spun at 4500 rpm for 5 minutes. The supernatant was removed and for each 25 ml of *E. coli* spun down, the cells were resuspended in 10 ml cold 50mM $CaCl_2$ and kept on ice for 30 minutes. The cells were spun again at 4500 rpm for 5 minutes and the $CaCl_2$ decanted. To each tube 1.5ml of fresh $CaCl_2$ was added and the cells resuspended and pooled into a single tube. To the cells 3.5 ml of sterile 50% glycerol was added. Aliquots containing 400 µl competent cells were prepared and stored at –80°C.

2.4.12.2 Transformation of competent *E. coli* by heat-shock

An aliquot of competent *E. coli* cells was thawed on ice. For each transformation, 100 µl was used and 1 µl of plasmid (between 10 pg – 100ng) or 10-20 µl of ligation product was added to the cells. A control mix, with no DNA was also prepared. The cells were incubated on ice for 30 minutes, at 42°C for 1 minute and then transferred back onto ice. To the heat-shocked cells, 1 ml of LB medium was added and they were incubated at 37°C under continuous shaking for 1 hour. After incubation, 100 µl and 200 µl of the cells were spread onto LB plates containing the appropriate antibiotic and incubated at 37°C overnight.

2.4.12.3 Isolation and analysis of *E. coli*

To check whether the *E. coli* transformants contained the expected plasmid either a colony PCR on putative transformants was performed, or a restriction digest test on the purified plasmid.

To perform the colony PCR, the colony picked by sterile pipette tip is used as the template DNA and is resuspended in the PCR mixture (2.4.1). A couple of colonies showing a

positive PCR result were picked using a sterile pipette tip, and used to inoculate 5-10 ml LB medium supplemented with the appropriate antibiotic, and grown at 37°C under continuous shaking overnight. Plasmid isolation (2.4.4) was performed and a glycerol stock was prepared (2.1.2.3.2).

To test the plasmid by performing a restriction digest test, individual colonies were picked using a sterile pipette tip, which were then used to inoculate 5-10 ml of LB medium supplemented with the appropriate antibiotic. These cultures were then grown in a continually shaking incubator set at 37°C overnight. Plasmid isolation (2.4.4) was performed to extract the plasmid. The plasmid was digested using restriction enzymes, and the product run on an agarose gel. A glycerol stock of a successful transformant was prepared.

2.4.13 Transformation of *Synechocystis*

The transformation method used was adapted from (Williams, 1988) by (Al-Haj, 2014). BG-11 liquid medium was inoculated with WT *Synechocystis* from a freshly restreaked plate. The cells were grown in an illuminated shaking incubator at 25°C at a light intensity of 40–100 μmol/m^2/s to an OD_{730} of 0.6–1.

X ml of the culture was taken and harvested by centrifugation at 3000x g for 10 minutes. X ml was calculated to ensure the final concentration of the cells was $4x10^8$ cells/ml before the DNA was added. To calculate initial concentration of the culture (cells/ml) the OD_{730} value was multiplied by $1.15x10^8$ cells/ml (Bishop, 2002). Cells were 'washed' with 2 ml fresh BG-11 medium and spun down again at 3000 xg for 5 minutes. The cell pellet was resuspended in 1ml BG-11 medium to give a final concentration of $4x10^8$ cells/ml.

For each transformation, 200 μl was aliquoted into 1.5 ml microfuge tubes and 0.4–2 μg of plasmid DNA was added. A 'no DNA' transformation was also prepared and acted as a control. The cells were incubated at 30°C in ~40 μmol/m^2/s for 4-6 hours. The cells were spread onto BG-11 agar plates and left to dry before incubation in a 30°C incubator for 2-3 days. The plates were then overlain with 3 ml of 0.6% of agar supplemented with the appropriate antibiotic and returned to the 30°C incubator. Colonies would typically appear after 7 days.

2.4.13.1 Isolation and analysis of *Synechocystis* transformants

Single colonies were picked using sterile toothpicks and streaked onto BG-11 agar plates supplemented with the appropriate antibiotic. Colonies were restreaked between 0 and 4 times unless otherwise mentioned. Genomic DNA from *Synechocystis* was extracted

(2.4.13.2), and used as the template in a PCR reaction in order to determine whether the transformant genome was homoplasmic and all copies of the genome had successfully integrated the DNA. If the PCR results indicated the transformant genome was heteroplasmic and there was a mixture of the original genome and the transformed genome, then further rounds of selection were carried out until homoplasmicity was reached by restreaking onto BG-11 agar containing the appropriate antibiotic. A glycerol stock (2.1.1.3.5) of successful transformants was prepared.

2.4.13.2 Genomic DNA extraction of *Synechocystis* for PCR analysis

To extract the genomic DNA the Chelex®-based method adapted from (Werner and Mergenhagen) was used. A loopful of cells from a plate was resuspended in 20 µl distilled water. To the cells, 20 µl of absolute ethanol was added, after 1 min, 200 µl of a 5% suspension of Chelex® 100 Resin (Bio-Rad) was added. The tubes were then vortexed briefly and boiled for 5 minutes. They were then cooled on ice and the cell debris pelleted with a 2 minutes spin at 13000 rpm. The supernatant containing the genomic DNA was transferred to a new tube, and 1 µl of supernatant used in a 25 µl PCR reaction.

2.4.14 The *codA*-5-FC method for the production of marker-less transformants

Strain 6803.IR706.codA (see section 5.2.3.2) from a freshly restreaked plate was used to inoculate 20-50 ml BG-11 medium. Cells were grown to an OD_{730} of 0.5-1.2, in an illuminated shaking incubator at 25°C at a light intensity of 40-100 µmol/m^2/s.

X ml of the liquid culture was harvested, washed and resuspended in 1 ml BG-11 medium, using the same method as in 2.4.13 to ensure the final concentration of the cells was $4x10^8$ cells/ml.

For each transformation, 10 µl was aliquoted into a 30 ml sterilin tube and 1 µg of DNA was added. A transformation mix where no DNA was added was also prepared to act as a control. The cells were then placed at 30°C in ~40 µmol/m^2/s with no shaking for 5 hours. To each tube, 2 ml of BG-11 was added and the cells were grown in an illuminated shaking incubator at 20°C at a light intensity of 20-40 µmol/m^2/s for 96 hours. 1ml and 0.5ml were then plated onto BG-11 agar plates supplemented with 0.5 mg/ml 5-FC. Colonies would appear after 7-12 days, and were picked and restreaked onto BG-11 agar plates supplemented with 0.5 mg/ml 5-FC. Genomic DNA was isolated and analysed to check for successful transformants.

2.4.14.1 Growth spot tests for analysing successful marker-less transformants

Growth spot tests were performed as a quicker method for selecting for successful transformants. A pipette tip was used to pick colonies from the BG-11 medium supplemented with 0.5 mg/ml 5-FC. The colony was then resuspended in 50 µl of BG-11, 10 µl was then aliquoted onto BG-11 and BG-11 supplemented with 200 µg/ml kanamycin. Once the liquid had dried, the plates were incubated at 30°C incubator in ~40 µmol/m^2/s light. Photos were taken of the plates after 3-7 days.

2.4.15 The *sacB*-sucrose method for the production of marker-less transformants

Strain 6803.IR706.sacB (see section 5.2.6.2) from a freshly streaked plate was used to inoculate 20 ml of BG-11 medium and the cells were grown to an OD_{730} of 0.8. The same protocol used to successfully obtain transformants using the *codA*-5-FC method was used (2.4.14) except 6803.IR706.sacB was used instead of 6803.IR706.codA and the transformation mix was plated onto BG-11 supplemented with 5% (w/v) sucrose.

2.5 Protein analysis

2.5.1 Isolation of crude whole protein extracts of *Synechocystis* strains

A flask containing 50 ml BG-11 was inoculated with the *Synechocystis* strain from a freshly restreaked plate and grown to an OD_{730} of 0.6-1.2. The cells were harvested by spinning the cells down at 3000 xg for 10 minutes. The cell pellet was then resuspended in X ml of Solution A, (0.8 M Tris-HCl pH 8.3, 0.2 M Sorbitol and 1% β-mercaptoethanol) where X ml is the OD_{730} value. This was done to obtain samples with approximately the same concentration. Samples were stored at -20°C for up to two months.

2.5.2 Glass beads used to isolate crude whole protein extracts of *Synechocystis* strains

A flask containing 50 ml BG-11 was inoculated with the *Synechocystis* strain from a freshly restreaked plate and grown to an OD_{730} of 0.6-1.2. The cells were harvested by spinning the cells down at 3000 xg for 10 minutes. The cell pellet was then resuspended in X ml of enzyme buffer A/B (see below), where X ml is the OD_{730} value, to obtain samples with the same concentration. 450 µl (800 µl for the enzyme assay) of the sample was then transferred to a 1.5 ml microfuge tube and ~100 µl of 212-300 µm acid washed glass beads (Sigma), were added to the cells. The sample was then treated to 5 cycles of vortexing for 2 minutes then placing on ice for 2 minutes. Cells were then spun at 3000 rpm at 4°C for 5 minutes. The supernatant was then spun down further at 13000 rpm at 4°C for 20 minutes to separate the soluble fraction and the insoluble fraction containing the membrane proteins.

Enzyme buffer A: 50 mM $MnCl_2$, 500 mM KCl, 1mM Tris/HCl (pH 7.5), 5 mM dithiothreitol (DTT), 0.05% (w/v) $NaHSO_3$ and 10% glycerol (Bohlmann et al., 1997).

Enzyme buffer B: 15 mM MOPS (pH 7.0), 10mM $MgCl_2$, 2 mM DTT (GE Healthcare), 1 x cOmplete™, EDTA-free Protease Inhibitor Cocktail tablet per 50 ml (Roche) (Entova, 2013).

2.5.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out using the Mini-Protean® Tetra System (Bio-Rad). Gels (80 x 83 x 1 mm gels) were prepared with a lower resolving gel, either 10 or 12% acrylamide, and a 5% acrylamide upper stacking gel (Table 2.7)

To prepare the crude protein extracts in solution A (2.5.1), 11 µl of 10% (w/v) SDS was added to a 100 µl aliquot of the crude protein. The samples were then boiled for 2 minutes and spun at 14000 rpm for 2 minutes to pellet cell debris.

To prepare samples from the crude protein extracts isolated using glass beads (2.5.2), 20 µl of 5 x Laemmli sample dye (see below) was added to an 80 µl aliquot of the crude protein sample along with 1 µl β-mercaptoethanol. Samples were then boiled at 99°C for 2 minutes, then spun at 13 000 rpm for 2 minutes.

Laemmli sample dye: 0.25 M, 40% (v/v) glycerol, 8% (w/v) SDS, 0.4% (w/v) Blue Bromophenol

For gels with 10 wells, 25 µl of each sample was loaded into each well. Gels were run in reservoir buffer (0.25 M Tris, 1.92 M Glycine and 1% SDS) at 120 V for 2 hours. PageRuler™ Prestained Protein Ladder (Thermo Scientific™) was loaded alongside samples in order to estimate protein size.

Table 2.7 Recipe for preparation of resolving and stacking gels.

	Resolving Gel (Volume for 1 gel)		**Stacking gel (Volume for 1 gel)**
	12%	**10%**	
H_2O	2.2 ml	2.45 ml	2.87 ml
40% Acrylamide/ bis-acrylamide solution, 37.5:1	1.5 ml	1.25 ml	0.5 ml
1.5 M Tris (pH 8.8)	1.3 ml	1.3 ml	0.5 ml
10% SDS	0.05 ml	0.05 ml	0.04 ml
10% Ammonium persulphate (AMPS)	0.05 ml	0.05 ml	0.04 ml
N, *N*, *N*′, *N*′-Tetramethylethane-1, 2-diamine (TEMED)	0.002 ml	0.002 ml	0.004 ml

2.5.4 Western blot analysis (Semi-dry)

After electrophoresis, the gel was soaked in Towbin buffer (Table 2.8), along with 6 pieces of 3 mm Whatman™ chromatography paper and an Amersham™ Hybond™-ECL nitrocellulose membrane (paper and membrane both cut to the same size as the gel). The blotter was set up in the following method, with 3 pieces of chromatography paper at the bottom, then the ECL nitrocellulose membrane, followed by the gel and the remaining 3 pieces of chromatography paper. The transfer was performed using a Bio-Rad Trans-Blot® SD Semi-Dry Transfer Cell at 20 V for 1 hour.

2.5.5 Immuno-detection

Once the proteins had been transferred onto the membrane, the membrane was blocked overnight in blocking buffer: 5% Skim milk powder in TBS (Table 2.8) at 4°C.

The milk was poured off and the membrane washed in TBS for 5-10 minutes. This was repeated three times, with gentle shaking. The membrane was then incubated with the primary antibody (Table 2.9) for 2-4 hours with gentle shaking at room temperature. After incubation with the primary antibody, the membrane was washed with TBS-T for 5-10

minutes with gentle shaking. This wash step was repeated three times. The membrane was then incubated with the secondary antibody (Table 2.9) in 5% Skim milk powder-TBS-T for 2 hours with gentle shaking. After incubation, the membrane was washed once more with TBS-T.

When horseradish peroxidase-linked secondary (ECL) antibodies were used, the membrane was incubated with SuperSignal® West Pico Chemiluminescence Substrate (Thermo Scientific™) for two minutes at room temperature. In a dark room, the membrane was exposed to a sheet of autoradiography Film – Amersham™ Hyperfilm™ ECL (GE Healthcare) and was developed using a Xograph automatic film developer. The time of exposure varied, depending on the strength of the signal.

For quantitative Western blot analysis, IRDye® secondary antibodies (Dylight™800) were used as the secondary antibody. Before detection the membrane was washed with TBS. Detection was carried out using the Odyssey® Infrared Imaging System (Li-COR Biosciences) and the infrared fluorescence signal was excited at 785 nm. Image Studio Lite 2 (Li-COR Biosciences) software was used to analyse the membrane and quantify the proteins.

Table 2.8 Buffer and solutions used for Western blot analysis

TBS	
Tris Base	20 mM
NaCl	137 mM
pH to 7.6 with HCl	1M
TBS with Tween-20 (TBS-T)	
TBS	As above
Tween-20	0.1% (v/v)
Towbin Buffer	
Tris	25mM
Glycine	192 mM
Methanol	20% (v/v)

2.5.6 Quantification of Western blot analysis

The IR fluorescent signals of the bands detected from the Western blot analysis were measured using the Odyssey® Infrared Imaging System (Li-COR Biosciences) and were analysed using Image Studio Lite 2, looking at the signals of the proteins from the 800 nm channel. The background value was defined by drawing a rectangle onto the membrane where no band was present. To quantify the signal, rectangles were drawn around each band to define the region for quantification; the background signal was automatically subtracted.

When normalising the target sample to the loading control, the infrared signal of the loading control was quantified. Each loading control value was divided by a single loading control, to give a proportional value. The target signal was then divided by their respective loading control to give a normalised IR fluorescent signal.

Antibodies

Table 2.9 shows a list of the primary and secondary antibodies used in this study.

Table 2.9 Antibodies used in this study for Western blot analysis.

Antibodies	Source	Dilution
Primary antibodies		
Anti-HA (polyclonal, produced in rabbit)	Sigma-Aldich	1:2000
Anti-GFP (polyclonal, produced in rabbit)	Thermo-Fisher Scientific™	1:2000
Anti-RbcL (polyclonal, produced in rabbit)	Gift from P. Nixon, Imperial College London, UK	1:8000
Anti-PsaD (polyclonal, produced in rabbit)	Gift from J. Gray, University of Cambridge, UK	1:1000
Secondary Antibodies		
ECL anti-Rabbit IgG, horseradish peroxidase-linked whole antibody (from donkey)	Amersham, GE Healthcare	1:5000
Anti-Rabbit IgG (H&L) DyLight™ 800 conjugated (from goat)	Thermo Scientific™	1:15000

2.6 Activity assays

2.6.1 Growth spot tests to assay expression of selectable markers

Cells either from a small loopful resuspended in 100 µl of BG-11 or from a liquid culture grown to an OD_{730} of 0.6-1 were used for growth spot tests. A 10-fold serial dilution was performed on the strains and 10 µl of each dilution was applied onto solid medium containing BG-11 supplemented with the appropriate selective agent. The plate was then left to dry and then incubated in the 30°C incubator in ~40 µmol/m^2/s light. The plates were inspected every day and photographs were taken after 3-7 days.

2.6.2 Induction of P_{nrsB} using Ni^{2+}

A 6.4 mM $NiCl_2.H_2O$ stock solution was prepared, and used for induction. Unless otherwise stated, the final concentration for induction was 6.4 µM. When incubating in liquid media, samples were induced for 4 hours, unless otherwise stated, before harvesting the samples for further analysis.

2.6.3 Detection of GFP using a luminescence spectrometer

To detect the emission of GFP from various *Synechocystis* strains expressing the protein, liquid cultures of the various strains were grown to a similar OD_{730}. For analysis 4ml was used in the PerkinElmer LS 55 Luminescence Spectrometer. The emission spectrum was recorded from 420 nm to 700 nm, with excitation at 395 nm and a scan speed of 500nm/min. The peak at 509 nm representing GFP was analysed.

In later studies using the luminescence spectrometer, 50 ml of BG-11 was inoculated from a freshly restreaked plate, to give a starting OD_{730}: ~0.05. Cells were then grown in low light, 20-40 µmol/m^2/s at 25°C, with shaking at 120 rpm until the OD_{730} reached ~0.6-0.8. X ml of the cultures were then harvested and resuspended in 4 ml of BG-11. X ml was calculated (see section 2.4.13) to ensure the final concentration of the cells was $4x10^8$ cells/ml.

2.6.4 GC-MS

2.6.4.1 Ethyl acetate extraction

The strains were inoculated in 25 ml of BG-11 and were grown to an OD_{730} of ~1.4. To perform the ethyl acetate extraction, 1 ml was centrifuged at 13000 rpm for 5 minutes. The supernatant was removed and transferred into another 1.5 ml microfuge tube. To the pellet and the supernatant, 500 µl ethyl acetate (HPLC grade) was added. The samples

were then heated at 50°C for 10 minutes before transferring the ethyl acetate layer to a GC vial. The vials were stored at -20°C.

2.6.4.2 Two-phase dodecane extraction

The strains were used to inoculate 50 ml of BG-11, which were then overlaid with 5 ml n-dodecane (dodecane) (99+%, Alfa Aesar). The liquid cultures were then incubated until they grew to an OD_{730} of ~1.2. The dodecane layer was then transferred into a GC vial and stored at -20°C.

2.6.4.3 GC-MS conditions and analysis

Limonene standards were prepared with analytical standard (S)-Limonene (Sigma Aldrich) in dodecane or ethyl acetate. All samples in dodecane were diluted 1:20 with ethyl acetate before loading onto the GC-MS

GC-MS was performed on an Agilent 7890A/5975C GC-MS. One µl of each sample was injected and separated using a HP-5MS column (35 m × 0.25mm; 0.25 µm film thickness). The analytical conditions were as follows: He (1.5 ml/min) as the carrier gas, splitless injection, and the oven program was set at 45°C for 1 min then 40°C/min to 320°C. Scanning method was used to obtain a library match and the SIM method was the used at 68 m/z. The internal standard used was 1-octanol.

2.6.5 Limonene - DPPH assay

The DPPH solutions were freshly prepared in dodecane before use. Spectroscopy was performed using the FLUOstar Omega microplate reader (BMG LABTECH) at 25°C. Standards were prepared by dissolving limonene in dodecane (100 µl, various concentrations) and were added to 100 µl of 0.2 mM DPPH into polypropylene 96-well plates (Grenier bio-one). For the enzyme assay (2.6.5.1), 100 µl of the dodecane overlay was added to 100 µl of 0.2 mM DPPH. A 200 µl aliquot of dodecane was used to blank the plate reader and replicates of each sample performed. Samples were shaken at 500 rpm before each cycle reading in orbital shaking mode and absorbance readings at 510 nm were taken every minute for 180 minutes. Reaction rates calculated ignored the first 10 minutes, as this region was noisy.

2.6.5.1 Enzyme assay

To prepare the enzyme for the enzyme assay, 105ml of BG-11 was inoculated from a freshly streaked plate and grown to an OD_{730}: ~1.1. Glass beads were used to isolate whole protein (2.5.1), and the samples were resuspended in 0.5 x X ml enzyme buffer B and 800

µl of resuspended cells. After using the glass beads, 40 µl from the upper fraction and 40 µl from the lower fraction were used for Western blot analysis, to confirm the presence of the enzyme.

A 100 µl aliquot of the upper fraction, containing the soluble proteins, was then used for the enzyme assay and was transferred to a GC vial, containing 300 µl of either 0 /1/ 1.5 mM GPP (Echelon Biosciences) in 15 mM MOPs, 10 mM $MgCl_2$, 2 mM DTT. A layer of dodecane of 250 µl was then overlaid on top of the cell extract with GPP. Samples were left at 25°C for 24 hours before vortexing the sample to terminate the assay. To test for the presence of limonene, a 100 µl sample was used in the FLUOstar Omega microplate reader with 100 µl of 0.2 mM DPPH (2.6.5).

2.7 *Synechocystis* toxicity tests

2.7.1 Toxicity tests with limonene

A liquid culture of *Synechocystis* was grown to an OD_{730} of ~0.8, unless otherwise stated, and 25 ml was added to each flask. Either (*R*)-(+)-limonene (97%, Sigma Aldrich) or (*S*)-(−)-limonene (96%, Sigma Aldrich) was added to individual flasks to a final concentration of 0, 0.02, 0.1, 0.2, 0.4 or 1% (v/v). The cultures were then grown at 25°C in 50 µmol/m²/s light. Photographs were taken after the addition of limonene after 0, 2, 4, 6 and 24 hours to record the viability of the culture.

2.7.1.1 Growth on solid media after 24 hours

24 hours after the addition of (*R*)-(+)-limonene, a 200 µl aliquot was spread onto BG-11 medium and grown at 30°C in ~40 µmol/m²/s for 4-6 hours. Photographs of the plates were taken after 7 and 16 days.

2.7.2 Toxicity test with limonene and dodecane

A liquid culture of *Synechocystis* was grown to an OD_{730} of 0.7, and 25 ml was aliquoted into 12 flasks. For six of the flasks 5 ml of n-dodecane (Alfa Aesar) was added to the cells. Different volumes of (*R*)-(+)-limonene (Sigma Aldrich) was added to the flasks with and without dodecane to give the following final concentrations of limonene: 0, 0.02, 0.1, 0.2, 0.4 and 1% (v/v). Photographs were taken at 0, 2, 18, 24, and then every 24 hours after the addition of limonene to record the viability of the culture.

2.7.3 Toxicity tests with styrene, *trans*-cinnamic acid and styrene oxide

To prepare the *trans*-Cinnamic acid (tCa), a stock of 100 g/L tCa (99+%, Alfa Aesar) was prepared in 95% ethanol. A liquid culture of BG-11 was inoculated with wild-type *Synechocystis* and was grown to an OD_{730} of either ~0.8 or ~0.1. For each flask, 25 ml of the liquid culture was added along with different volumes of styrene (99%, Alfa Aesar), tCa and styrene oxide (98+%, Alfa Aesar). A culture without styrene, tCa or styrene oxide present acted as the control. Photographs were taken immediately after the addition of the three chemicals and twice a day thereafter to record the viability of the cells.

2.8 Construction of two-gene operons

2.8.1 Construction of two-gene operon containing *ble* and *gfp*

The PCR product, Ble_IG2 (0.4kb), was amplified via PCR using Ble_Fw and Ble_IG2_Re. The PCR product, IG2_GFP (0.7 kb), was also amplified via PCR, using GFP_Re and GFP_IG2_Fw. A PCR was then performed using Ble_IG2 and IG2_GFP as the template with Ble_Fw and GFP_Re to create, ble_IG2_GFP (1.1 kb). The same PCR reactions were performed with different primers (Table 2.10) to create Ble_IG4_GFP, Ble_IG5_GFP, Ble_IG0bp_GFP and Ble_IG1bp_GFP (all 1.1 kb).

Table 2.10 Primers used to create the PCR products of the two-gene operons

Template	Forward Primer	Reverse Primer	PCR product
pZΔES	Ble_Fw	Ble_IG2_Re	ble_IG2
4op	GFP_IG2_Fw	GFP_Re	IG2_GFP
ble_IG2 and IG2_GFP	Ble_Fw	GFP_Re	ble_IG2_GFP
pZΔES	Ble_Fw	Ble_IG4_Re	ble_IG4
4op	GFP_IG4_Fw	GFP_Re	IG4_GFP
ble_IG4 and IG4_GFP	Ble_Fw	GFP_Re	ble_IG4_GFP
pZΔES	Ble_Fw	Ble_IG5_Re	ble_IG5
4op	GFP_IG5_Fw	GFP_Re	IG5_GFP
ble_IG5 and IG5_GFP	Ble_Fw	GFP_Re	ble_IG5_GFP
pZΔES	Ble_Fw	Ble_IG0bp_Re	ble_IG0bp
4op	GFP_IG0bp_Fw	GFP_Re	IG0bp_GFP
ble_IG0bp and IG0bp_GFP	Ble_Fw	GFP_Re	ble_IG0bp_GFP
pZΔES	Ble_Fw	Ble_IG1bp_Re	ble_IG1bp
4op	GFP_IG1bp_Fw	GFP_Re	IG1bp_GFP
ble_IG1bp and IG1bp_GFP	Ble_Fw	GFP_Re	ble_IG1bp_GFP

Each Ble_IGX_GFP PCR fragment was then digested with AseI and BglII and cloned into pLAH.cpc between the NdeI and BamHI site. The insertion of Ble_IGX_GFP into pLAH.cpc to produce various pLAH.cpc.ble.IGX.GFP expression plasmids was confirmed by PCR using Ben.ups.F and C.Frag.BR. Sequencing with primers Pcpc_ins_F, ble_Fw and Out_GFP_R also confirmed the results of the PCR.

2.8.2 Construction of two-gene operon containing *ble* and *splB*

Primers used to transcriptionally couple *splB* and *gfpuv* were designed in a similar manner to those used to amplify *ble* when the intergenic region was zero base and one base (Table 2.11).

Table 2.11 Table containing the primers used to create the PCR products containing the two-gene operon with *splB* and *gfpuv*.

Template	Forward Primer	Reverse Primer	PCR product
pASapI-SplB	splB_F	splB_IG0bp_R	splB_IG0bp
4op	GFP_0bp_F	GFP_Re	IG0bp_GFP (splB)
splB_IG0bp and IG0bp_GFP (splB)	splB_F	GFP_Re	splB_IG0bp_GFP
pASapI-SplB	splB_F	splB_IG1bp_R	splB_IG1bp
4op	GFP_1bp_F	GFP_Re	IG1bp_GFP (splB)
splB_IG1bp and IG1bp_GFP (splB)	splB_F	GFP_Re	splB_IG1bp_GFP

The PCR products, splB_IG0bp_GFP and splB_IG1bp_GFP, were digested with AseI and BglII, and cloned into pLAH.cpc at the NdeI and BamHI restriction sites to produce pLAH.splB.0bp.GFP and pLAH.splB.1bp.GFP. Construction of these plasmids was confirmed by performing a test digest using EcoRV and NdeI. Sequencing with primers pcpc_ins_fw and GFP_Fw_0bp/GFP_Fw_1bp confirmed the presence of the insert in pLAH.splB.IG0bp.GFP and pLAH.splB.IG1bp.GFP.

2.9 Bioinformatics

Alignment and sequence analysis was performed using MacVector 12.6.0. Table 2.12 list the following free online tools also used in this study.

Table 2.12 List of databases and websites used in this study.

Protein and DNA sequences Database	
Sequence for *Synechocystis* genome	http://genome.microbedb.jp/cyanobase/Synechocystis
NCBI	http://www.ncbi.nlm.nih.gov/
Bioinformatics tools	
Primer design	http://www.basic.northwestern.edu/biotools/oligocalc.html
Reverse complement	http://www.bioinformatics.org/sms/rev_comp.html
Sequence alignment	http://www.ebi.ac.uk/Tools/msa/clustalw2/
Restriction digest	http://nc2.neb.com/NEBcutter2/
Calculate MW of Proteins	http://web.expasy.org/protparam/
Ligation calculator	http://www.insilico.uni-duesseldorf.de/Lig_Input.html
Intergenic sequences search	http://meta.microbesonline.org/operons/gnc1148.html
Translation tool	http://www.attotron.com/cybertory/analysis/trans.htm
BPROM	http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb
BLAST ®	http://blast.ncbi.nlm.nih.gov/Blast.cgi
Operon Predictor	http://www.microbesonline.org/operons/gnc1148.html
Codon optimisation for SO operon for *Synechocystis*	https://www.idtdna.com/CodonOpt

Chapter 3 Expression of limonene synthase in *Synechocystis*

3.1 Introduction

3.1.1 Limonene

Limonene is a C_{10} cyclic monoterpene that is produced as an essential oil in plants. Many plants naturally produce limonene as a pollinator attractant and a defence mechanism against herbivores (Maruyama et al., 2001).

Recently there has been growing interest in limonene and its derivatives as they can be used in a wide range of applications (Figure 3.1). Limonene could be used as an alternative aviation fuel as it was demonstrated to have properties (e.g. freezing point, volatility and viscosity) that were comparable to Jet A-1, one of the most commonly used aviation fuels in the world (Chuck and Donnelly, 2014). Limonene in the hydrogenated form also has favourable properties for use as a jet fuel (e.g. high energy density, immiscible in water and a low freezing point) (Alonso-Gutierrez et al., 2013) and as a diesel fuel additive to enhance cold weather performance (Tracy et al., 2009). Limonene has a pleasant smell, and is often used as a fragrance in perfumes and cosmetics and is present in cleaning products. The oxygenated forms such as carvone and perillyl alcohol are also useful chemicals with applications in the pharmaceutical, food and cosmetic industries (Alonso-Gutierrez et al., 2013; Carter et al., 2003).

Limonene is usually obtained as a by-product from the orange juice industry, however the process of extracting limonene (mechanical pressurising and distillation) is energy intensive and costly, making it unsustainable for biofuel production (Halfmann et al., 2014b). Both the d- and l- enantiomers of limonene exist, due to the presence of a chiral centre, and have different properties and uses. The limonene produced from citrus fruits is often 70-98% of the d-enantiomer (also called (+)- and *R*-limonene) which has a citrus smell and is the enantiomer used as a flavour/fragrance ingredient (Jongedijk et al., 2016). The l-limonene (also called (−)-limonene and *S*-limonene) has a piney, turpentine odour. Both (−)-perillyl alcohol and (−)-carvone are produced from the l-limonene enantiomer. Limonene is sold at 9-10 $/kg, whereas the derivatives of limonene are sold at a higher price; (−)-menthol costs 15-20 $/kg and (−)-carvone costs 15-20 $/kg (Lange, 2015).

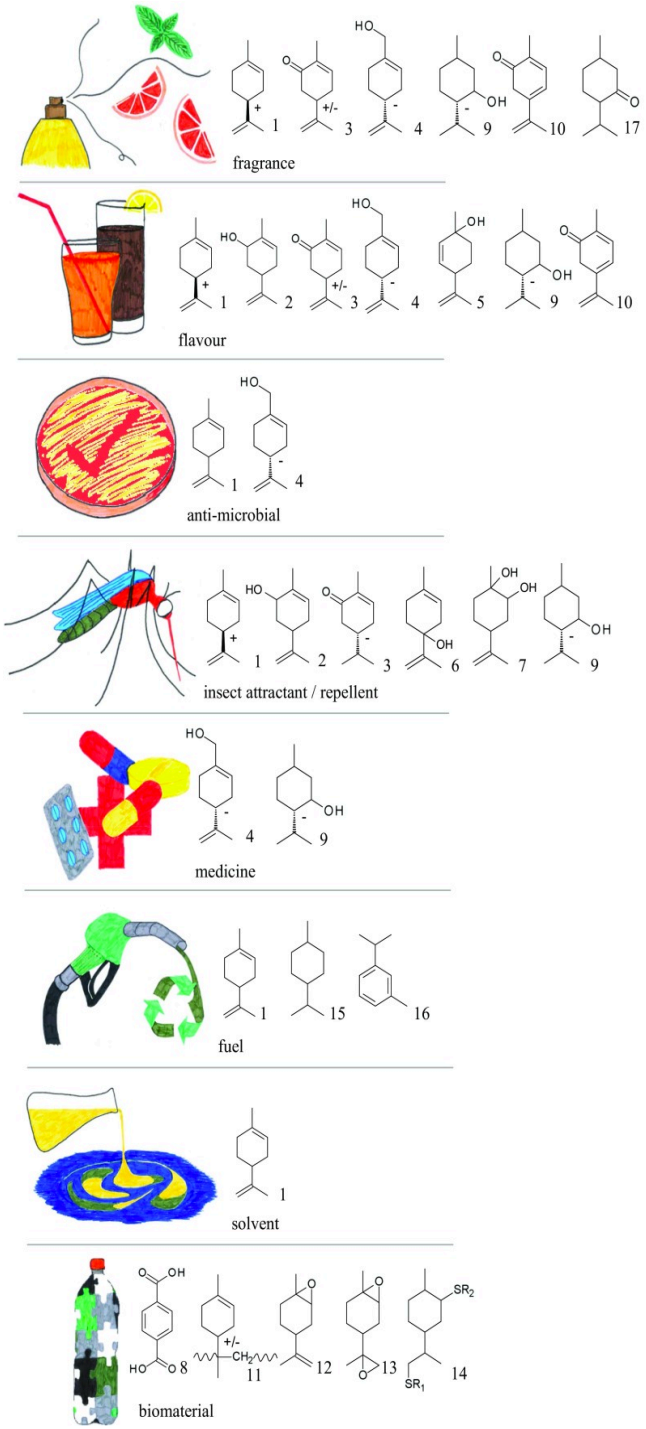


Figure 3.1 Applications of limonene and derivatives of limonene

1. Limonene; 2. carveol; 3. carvone; 4. perillyl alcohol; 5. p-mentha-2,8-diene-1-ol; 6. p-mentha-1,8-dien-4-ol; 7. p-menth-8-ene-1,2-diol; 8. terephthalic acid; 9. menthol; 10. dehydrocarvone, 11. polylimonene; 12. limonene monoepoxide; 13. limonene di-epoxide; 14. product of thiol di-addition (R1 and R2 are thiol-side groups e.g., 2-mercaptoethanol, methyl thioglycolate, or thioglycerol; 15. 1-isopropyl-4-methylcyclohexane; 16. m-cymene; 17. menthone. + and – indicate where a single enantiomer is used; +/– means either enantiomer can be used, but not as a mixture. Figure from (Jongedijk et al., 2016).

In plants, limonene and other monoterpenes are produced from the common precursor geranyl diphosphate (GPP). Two isoprene units DMAPP and IPP are produced from the MEP pathway or the MVA pathway and are combined by geranyl diphosphate synthase (GPPS) to produce GPP (Figure 1.7). Limonene is then produced from GPP using limonene synthase (LS). Some plant species produce one enantiomer, whereas others produce both in varying ratios (Maruyama et al., 2001).

3.1.2 Limonene production in microorganisms

Successful attempts have been made to produce limonene in *E. coli* and *S. cerevisiae* (Alonso-Gutierrez et al., 2013; Behrendorff et al., 2013; Carter et al., 2003; Jongedijk et al., 2015; Willrodt et al., 2014). In *E. coli* limonene production from sugar was achieved after expressing GPPS from Grand fir (*Abies grandis*) and LS from spearmint (*Mentha spicata*) (Carter et al., 2003). The yield of limonene production was further improved by expressing heterologous genes from *S. cerevisiae* to introduce the MVA to work alongside the MEP pathway and produce more IPP and DMAPP (Alonso-Gutierrez et al., 2013; Willrodt et al., 2014). The production of limonene has also been achieved in *S. cerevisiae* by expressing LS from lemons (*Citrus limon*), beefsteak plants (*Perilla frusctascens*) and spearmint (Behrendorff et al., 2013; Jongedijk et al., 2015).

3.1.3 Aims and objectives

At the time of starting the project, the production of limonene in cyanobacteria had not been achieved. The main aim of the work in this chapter was to attempt the synthesis of limonene in *Synechocystis* through genetic engineering, and to increase the production of limonene by making further genetic modifications to the metabolic pathway. As GPP (the substrate for LS) is produced from the MEP pathway in *Synechocystis*, the production of limonene was attempted by expressing a LS gene under various promoters (Figure 3.2). Alongside this work, attempts were made to increase the pool of GPP for LS, by knocking out pathways downstream of GPP. The toxicity of limonene towards *Synechocystis* was also investigated, as the toxicity of limonene towards other microorganisms has been demonstrated.

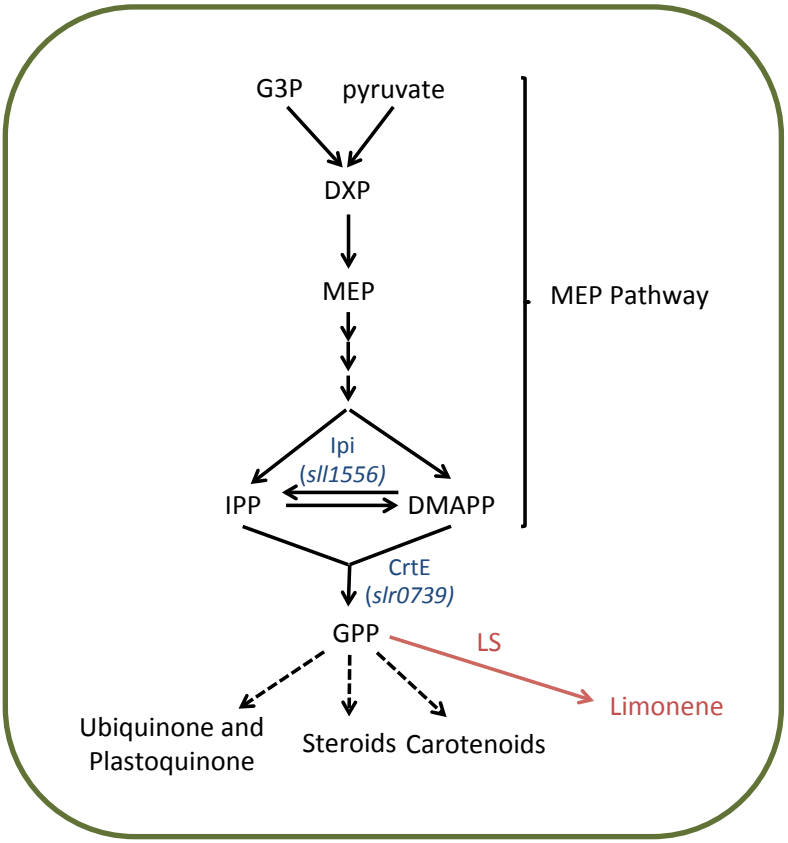


Figure 3.2 Native MEP pathway for isoprenoid synthesis and engineered limonene pathway in *Synechocystis*

Schematic showing the methyerythritol phosphate (MEP) pathway. Native *Synechocystis* genes in blue and non-native pathway in red. Abbreviations used: G3P = glyceraldehyde 3-phosphate, DXP = 1-deoxy-d-xylulose 5-phosphate, MEP = methylerythritol-4-phosphate, IPP = isopentenyl diphosphate, DMAPP = dimethylallyl diphosphate, GPP = geranyl diphosphate, LS = limonene synthase.

3.2 Results

3.2.1 Toxicity of limonene towards *Synechocystis*

One of the main problems with producing novel compounds in cyanobacteria is the problem with toxicity. Monoterpenes such as limonene are thought to interfere with the cell membrane (Brennan et al., 2012). The external addition of limonene proved to be toxic, at low concentrations, for *E. coli* (0.025% (v/v)) and *S. cerevisiae* (~0.0075% (v/v)), (Brennan et al., 2012; Dunlop et al., 2011). Although *E. coli* has been engineered to produce limonene, the levels of production were not yet toxic to the cells (Dunlop et al., 2011). Therefore toxicity tests were performed by adding limonene exogenously to cultures of *Synechocystis* to determine the potential effects of introducing a limonene production pathway into *Synechocystis.*

3.2.1.1 Initial toxicity tests with limonene

An initial toxicity test was performed on WT *Synechocystis* with different amounts of (*R*)-limonene added to the cultures (Figure 3.3A). After two hours the cultures supplemented with limonene appeared unhealthy in comparison to the control strain that had no (*R*)-limonene added. After 24 hours, the same cultures appeared colourless, which suggested the cells had died, so a sample of each culture was taken and grown onto solid BG-11 medium (Figure 3.3B). After 16 days, no colonies were observed on the plates that had 1% or 0.4% (v/v) of limonene added to the cells; colonies were present on some of the plates when 0.02%, 0.1% and 0.2% were added to the cultures, with fewer colonies obtained as the concentration increases. These results confirm that limonene is toxic towards *Synechocystis* at concentrations as low as 0.02% (v/v), and therefore production of the compound in the cyanobacterium should ideally be inducible (see section 3.2.2 below) so that biomass can be produced prior to inducing synthesis of limonene.

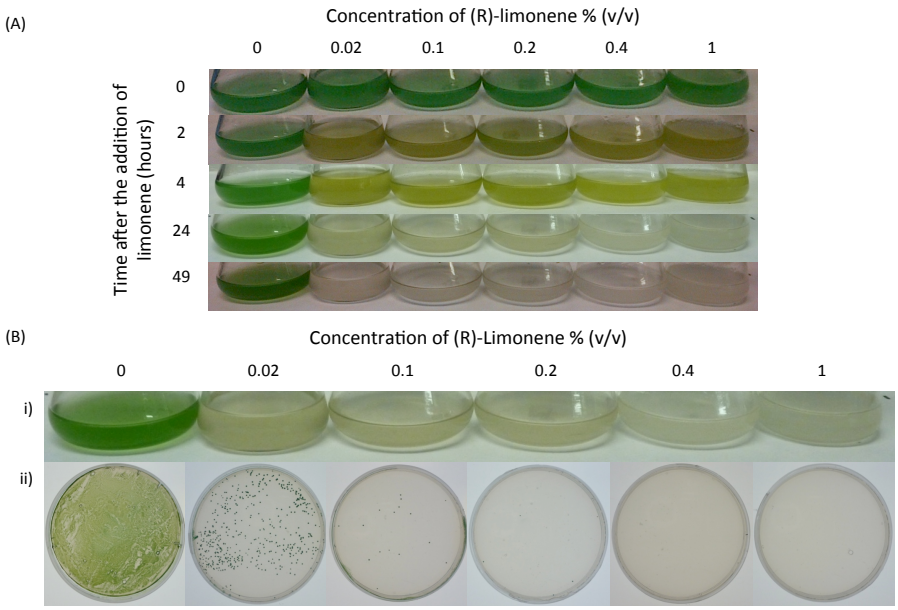


Figure 3.3 Photographs of toxicity tests taken after the addition of (*R*)-limonene to a culture of WT *Synechocystis* with a starting OD_{730} of 0.83.

(A) Photographs of the cultures were taken after 2, 4, 14 and 49 hours after the addition of (*R*)-Limonene. (B) i) culture at 24 hours, grown in BG-11 medium ii) plating on solid BG-11, photograph taken after 16 days.

3.2.1.2 Mitigate toxicity of limonene with dodecane

From the previous results, limonene appears to be toxic towards *Synechocystis*. Two-phase culture systems with various solvents have been effective in reducing the toxicity of limonene towards *S. cerevisiae* (Brennan et al., 2012). Dodecane has been used as the solvent in a two-phase system and was effective in reducing the toxicity of geraniol and farnesene towards *Synechocystis* (Al-Haj, 2014).

As before, different concentrations of (*R*)-limonene were added to a culture of *Synechocystis,* but in this case with and without a layer of 20% (v/v) dodecane to determine whether this immiscible hydrocarbon layer could mitigate the toxic effect of limonene by sequestering it (Figure 3.4). Similar results to those seen in section 3.2.1.1 were obtained in the absence of dodecane after 48 hours, and limonene was confirmed to be toxic at concentrations as low as 0.02% (v/v) limonene. However, in the initial tests, the cultures were only observed for 49 hours, whereas in this case the culture was observed for a longer period of time and after 72 hours, the culture of *Synechocystis* in the presence of 0.02% (v/v) limonene appeared slightly green and started to recover. This recovery is also seen after 228 hours when 0.1% (v/v) limonene was added to the cells. This may be due the degradation or loss of the volatile limonene from the medium, and recovery of a few 'survivor' cells, or the adaptation of a few cells to the limonene.

As shown in Figure 3.4, growth of *Synechocystis* is unaffected by an overlay of dodecane. More significantly, the dodecane allowed growth in the presence of limonene, even at the highest concentration tested – 1% (v/v). This illustrates that the toxic effect of limonene build up in the medium could be mitigated by the use of an overlay. Furthermore, if we assume that limonene can be excreted from engineered cells then the removal of the compound in the media into the overlay could extend the lifetime of a limonene-producing culture (Alonso-Gutierrez et al., 2013).

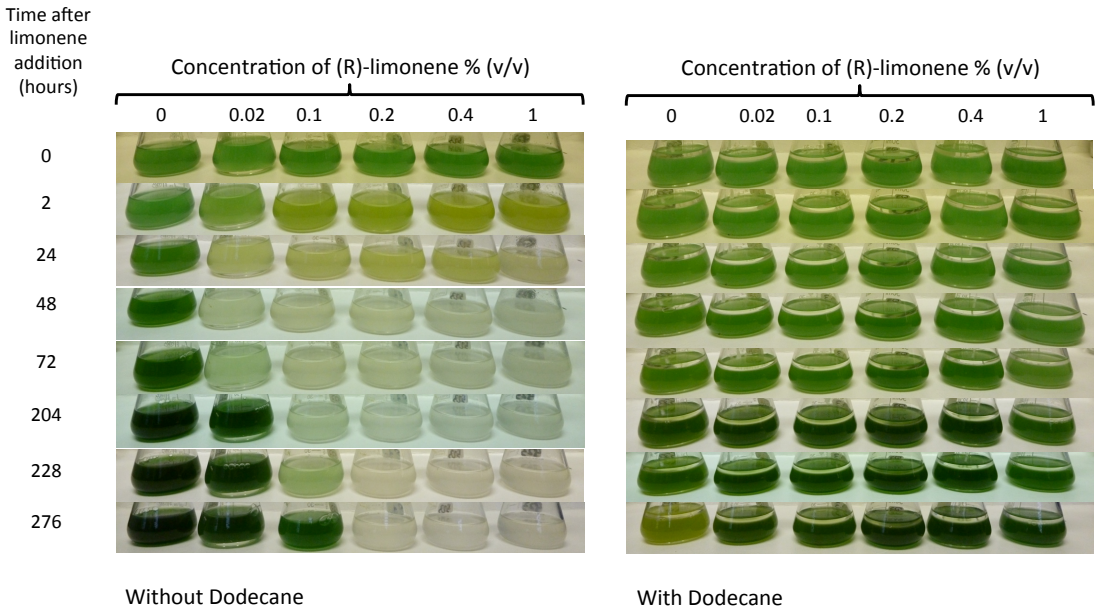


Figure 3.4 Toxicity tests following the addition of (*R*)-limonene to a culture of WT *Synechocystis* with a starting OD_{730} of 0.7, without and with a 20% (v/v) overlay of dodecane.

Photographs taken after 2 hours, then every 24 hours. (Note: not all photographs are shown).

3.2.1.3 Toxicity test with (S)-limonene

The toxicity of the (*S*)-limonene enantiomer was also tested to determine the effect on *Synechocystis* (Figure 3.5). (*S*)-Limonene was added to a culture of *Synechocystis* grown to OD_{730}: 0.5 and showed similar results to the addition of (*R*)-limonene. The cells are affected by the presence of 0.02% (v/v) limonene after 24 hours, but the cells recover at 0.02% (v/v) after 72 hours, which was also seen with (*R*)-limonene.

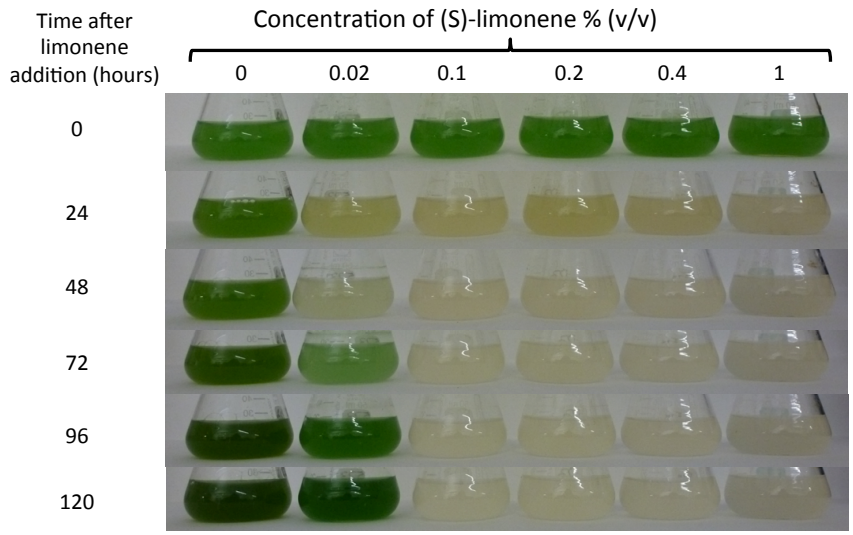


Figure 3.5 Toxicity tests after the addition of (*S*)-limonene to a culture of WT *Synechocystis* with a starting OD_{730}: 0.4.3. Photographs taken every 24 hours.

3.2.2 Expression of limonene synthase in *Synechocystis*

Interest in cyanobacteria has grown over the past few years, due to their ability to produce novel compounds directly from CO_2 using light energy and to grow on non-arable land (Machado and Atsumi, 2012; Ruffing, 2011). Limonene and its derivatives have a number of applications, including potential use in jet fuels (Jongedijk et al., 2016). Limonene production has not yet been attempted in cyanobacteria, although the substrate for LS, GPP, is produced. Therefore, the production of limonene was attempted in *Synechocystis* by the heterologous introduction of LS.

3.2.2.1 Design of a synthetic limonene synthase gene

In order to select a suitable LS sequence for gene design, the literature was searched for well characterised enzymes. LS was found to be characterised from a number of plants such as *Perilla frutescens*, *Schizonepeta tenuifolia*, *Agastache rugosa*, *Citrus limon*, *Abies grandis* and *Mentha spicata* (Bohlmann et al., 1997; Colby et al., 1993; Lücker et al., 2002; Maruyama et al., 2001, 2002; Yuba et al., 1996). The (-)-4S-Limonene synthase found in spearmint (*Mentha spicata*) (Accession number: L13459), which principally produces the *S*-limonene, was selected as it has been successfully expressed in *E. coli* for the production of limonene and has been characterised in more detail than the other LSs.

Other researchers had reported that expression of the full-length LS led to the formation of inclusion bodies in *E. coli*, but a truncated version of LS from *Mentha spicata* lacking an N-terminal transit sequence was shown to be functional in *E. coli* and *S. cerevisiae* (Behrendorff et al., 2013; Williams et al., 1998) The first 57 amino acids up to the

conserved RRX_8W motif were removed, as they make up the cleavable transit peptide for plastid localisation in plants, and therefore were not required for expression in cyanobacteria (Lücker et al., 2002). The threonine and glutamate residues just before RRX_8G were retained to provide stability to the enzyme (Varshavsky, 1997). For our gene design, the C-terminal end of the various LSs were compared, before a HA-epitope sequence (-YPYDVPDYA*) was added to enable the detection of LS (Figure 3.6). The LS sequence gene was codon-optimised for *Synechocystis* and designed with an NdeI at the 5' end of the gene and a BamHI site at the 3' end of the gene. The NdeI start site also provides the start codon for the truncated LS gene. The designed gene, *ls*, was synthesised by GeneArt and provided on plasmid ls_pMA-RQ (See Appendix).

```
Ms_LS     1 MALKVLSVATQMAIPSNLTTCLQPSHFKSSPKLLSSTNSSSRSRLRVYCSSSQLTTERRS  60
ls        1 ------------------------------------------------------MTERRS   6
                                                                   *****

Ms_LS    61 GNYNPSRWDVNFIQSLLSDYKEDKHVIRASELVTLVKMELEKETDQIRQLELIDDLQRMG 120
ls        7 GNYNPSRWDVNFIQSLLSDYKEDKHVIRASELVTLVKMELEKETDQIRQLELIDDLQRMG  66
             ************************************************************

Ms_LS   121 LSDHFQNEFKEILSSIYLDHHYYKNPFPKEERDLYSTSLAFRLLREHGFQVAQEVFDSFK 180
ls       67 LSDHFQNEFKEILSSIYLDHHYYKNPFPKEERDLYSTSLAFRLLREHGFQVAQEVFDSFK 126
             ************************************************************

Ms_LS   181 NEEGEFKESLSDDTRGLLQLYEASFLLTEGETTLESAREFATKFLEEKVNEGGVDGDLLT 240
ls      127 NEEGEFKESLSDDTRGLLQLYEASFLLTEGETTLESAREFATKFLEEKVNEGGVDGDLLT 186
             ************************************************************

Ms_LS   241 RIAYSLDIPLHWRIKRPNAPVWIEWYRKRPDMNPVVLELAILDLNIVQAQFQEELKESFR 300
ls      187 RIAYSLDIPLHWRIKRPNAPVWIEWYRKRPDMNPVVLELAILDLNIVQAQFQEELKESFR 246
             ************************************************************

Ms_LS   301 WWRNTGFVEKLPFARDRLVECYFWNTGIIEPRQHASARIMMGKVNALITVIDDIYDVYGT 360
ls      247 WWRNTGFVEKLPFARDRLVECYFWNTGIIEPRQHASARIMMGKVNALITVIDDIYDVYGT 306
             ************************************************************

Ms_LS   361 LEELEQFTDLIRRWDINSIDQLPDYMQLCFLALNNFVDDTSYDVMKEKGVNVIPYLRQSW 420
ls      307 LEELEQFTDLIRRWDINSIDQLPDYMQLCFLALNNFVDDTSYDVMKEKGVNVIPYLRQSW 366
             ************************************************************

Ms_LS   421 VDLADKYMVEARWFYGGHKPSLEEYLENSWQSISGPCMLTHIFFRVTDSFTKETVDSLYK 480
ls      367 VDLADKYMVEARWFYGGHKPSLEEYLENSWQSISGPCMLTHIFFRVTDSFTKETVDSLYK 426
             ************************************************************

Ms_LS   481 YHDLVRWSSFVLRLADDLGTSVEEVSRGDVPKSLQCYMSDYNASEAEARKHVKWLIAEVW 540
ls      427 YHDLVRWSSFVLRLADDLGTSVEEVSRGDVPKSLQCYMSDYNASEAEARKHVKWLIAEVW 486
             ************************************************************

Ms_LS   541 KKMNAERVSKDSPFGKDFIGCAVDLGRMAQLMYHNGDGHGTQHPIIHQQMTRTLFEPFA- 599
ls      487 KKMNAERVSKDSPFGKDFIGCAVDLGRMAQLMYHNGDGHGTQHPIIHQQMTRTLFEPFAY 546
             ************************************************************

Ms_LS   600 -------- 599
ls      547 PYDVPDYA 554
```

Figure 3.6 Alignment of the full length *Mentha spicata* limonene synthase (Ms_LS) aligned with the modified limonene synthase (ls) truncated at the N-terminus and a HA-tag at the C-terminus.

The first 55 amino acids are removed from the full-length LS sequence.

3.2.2.2 Creation of a transgenic strain of *Synechocystis* containing limonene synthase under the control of the nickel inducible promoter

As limonene was shown to be very toxic towards *Synechocystis*, the expression of LS under the control of the nickel inducible promoter, P_{nrsB}, was initially attempted. To create the transgenic strains of *Synechocystis* expressing LS under the control of P_{nrsB}, the expression plasmid pLAH.nrsB.LS was created to introduce *ls* at the *psbAII* site (Figure 3.7). The *psbAII* gene encodes the D1 protein in PSII, and deletion of this gene does not affect growth as the *psbAIII* gene expresses the D1 protein in its absence (Mohamed et al., 1993) .

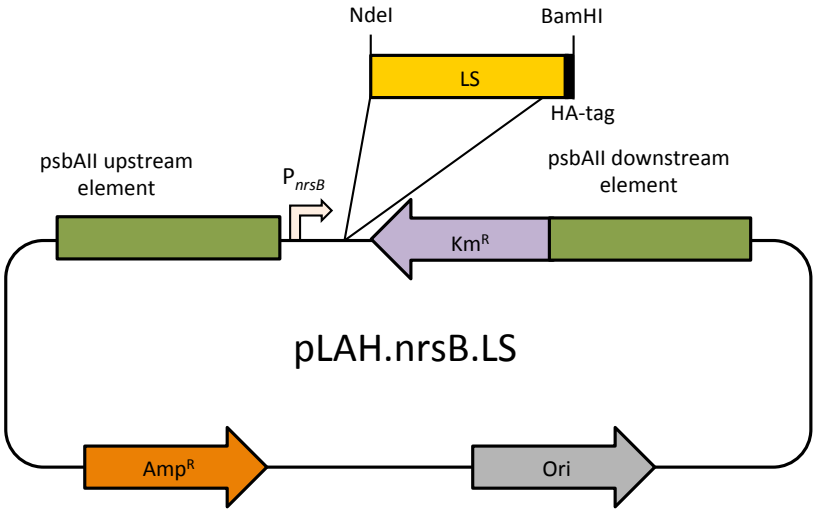


Figure 3.7 Diagram of the expression plasmid pLAH.nrsB.LS

The *ls* gene digested from ls_pMA-RQ was cloned into pLAH.nrsB to create pLAH.nrsB.LS, to enable the integration of *ls* under the control of P_{nrsB} at the *psbAII* site.

Using NdeI and BamHI, *ls* was digested from ls_pMA-RQ and cloned into the NdeI and BamHI sites of pLAH.nrsB to create pLAH.nrsB.LS. A restriction digest with NdeI and sequencing analysis confirmed the insertion of *ls* into pLAH.nrsB.

WT *Synechocystis* was transformed with pLAH.nrsB, and after selection for kanamycin resistance, colonies appeared after seven days. Eight colonies were picked after 8 rounds of selection on kanamycin and PCR analysis was performed to screen for putative transformants of 6803.nrsB.LS. Three clones contained the *ls* gene at the *psbAII* locus confirming the successful creation of 6803.nrsB.LS strains. As there are no copies of the WT genome detected by the PCR analysis, all strains were considered to be homoplasmic for the engineered genome (Figure 3.8).

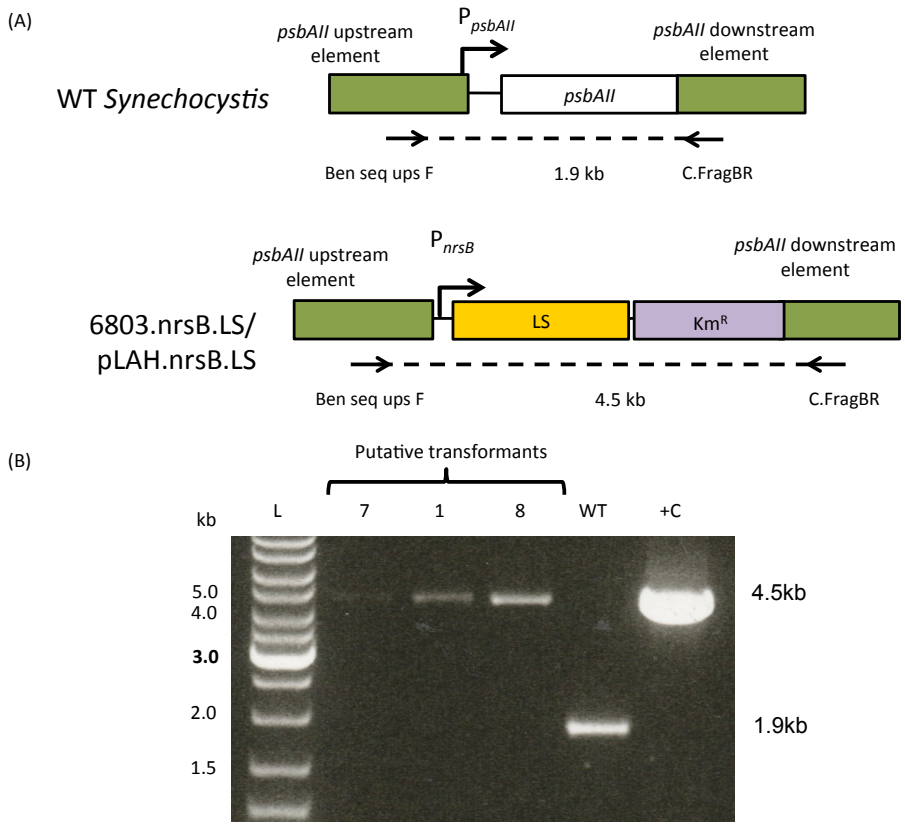


Figure 3.8 PCR confirmation of successful transformants of *Synechocystis* containing *ls* under the control of P_{nrsB}.

(A) Schematic showing the binding sites and expected fragments sizes for WT *Synechocystis*, pLAH.nrsB.LS expression plasmid and successful 6803.nrsB.LS transformants. (B) PCR screening of putative transformants. WT *Synechocystis* (WT) and pLAH.nrsB.LS (+C) were used as controls. The single band at 4.5 kb confirmed homoplasmic strains were produced.

3.2.2.3 Expression of limonene synthase under the control of the nickel promoter

LS was designed with an HA-epitope tag at the C-terminus to enable detection of expression in *Synechocystis* by using commercially available anti-HA antibodies.

Three transformants of 6803.nrsB.LS were grown to an OD_{730}: 0.8-1 before induction with 6.4 µM Ni^{2+} for 4 hours. The Ni^{2+} concentration was chosen as it has been used previously to induce P_{nrsB} and is thought to induce the maximal response in P_{nrsB} (Peca et al., 2008). Western blot analysis was performed with anti-HA antibodies on the crude extract of these transformants on 6803.nrsB.LS using horseradish peroxidase-linked secondary antibodies to measure the chemiluminescence on photographic film. The expression of LS was detected in all three transformants of 6803.nrsB.LS when induced with Ni^{2+} but in the absence of Ni^{2+} the expression of LS was not detected (Figure 3.9). These results confirm the creation of a strain of *Synechocystis*, 6803.nrsB.LS, which expresses LS under the P_{nrsB}.

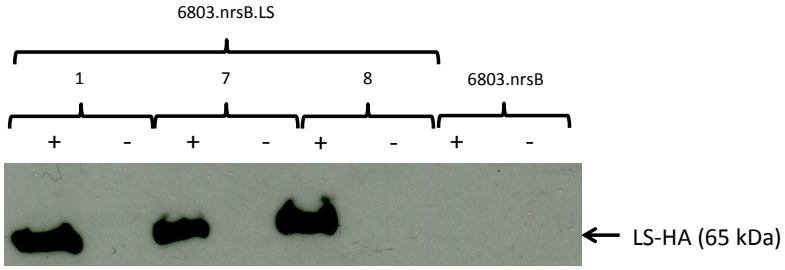


Figure 3.9 Western blot analysis showing the presence of HA-tagged LS in successful 6803.nrsB.LS transformants that have been induced with Ni^{2+}.

Western blot analysis was performed on samples of 6803.nrsB.LS that were either uninduced (−) or induced (+) with 6.4 μM Ni^{2+} for 4 hours. The ECL method to detect recombinant LS with anti-HA antibodies and ECL anti-rabbit IgG antibodies. Transformant line 6803.nrsB was used as a negative control. The expected band size of LS (65 kDa) was seen in the 6803.nrsB.LS transformants following induction with Ni^{2+}.

3.2.2.3.1 Investigating different induction conditions

Western blot analysis using anti-HA antibodies was performed on a single transformant of 6803.nrsB.LS_8 using the Odyssey® Infrared Imaging System, to investigate the effect of the length of induction and the concentration of Ni^{2+} used on the expression of LS (Figure 3.10). To test the length of induction a culture of 6803.nrsB.LS_8 was grown to an OD_{730}: 0.8 and divided into four flasks. Each flask induced with 6.4 μM Ni^{2+} and cells were harvested after 2, 4, 8 and 24 hours. By increasing the length of induction the expression of LS was seen to increase.

To test the effect of the concentration of nickel, a culture of 6803.nrsB.LS_8 was grown to an OD_{730}: 0.8, and divided into five flasks, with different concentrations of Ni^{2+} used for induction. The expression of LS with 0.4 μM Ni^{2+} was weak in comparison to the use of 3.2, 6.4, 12.8, and 19.2 μM Ni^{2+}. The expression of LS, when induced with higher concentrations of Ni^{2+} (12.8, and 19.2 μM), did not appear to be significantly greater than when 3.2 and 6.4 μM Ni^{2+} are used. This suggests that increasing the Ni^{2+} beyond 3.2 μM had limited influence on the expression level.

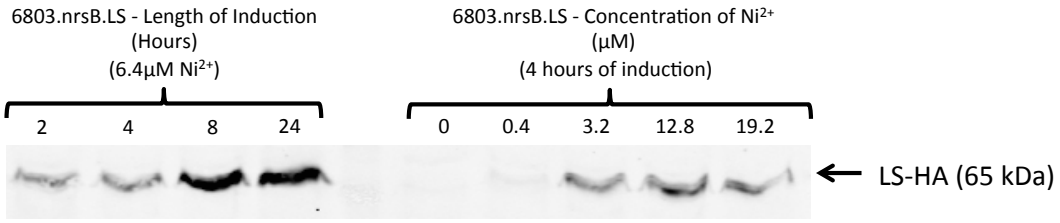


Figure 3.10 Western blot analysis comparing the expression of HA-tagged LS in 6803.nrsB.LS transformants induced with different times of induction and different concentrations of Ni^{2+}.

Western blot analysis was performed with whole cell extracts with anti-HA antibodies using the Odyssey® Infrared Imaging System for detection. Samples were run on a 12% SDS-PAGE gel. Protein size was determined using the PageRuler™ Prestained Protein ladder (Thermo Scientific). Samples of 6803.nrsB.LS were induced with 6.4 μM Ni^{2+} for 2, 4, 8 and 24 hours and samples of 6803.nrsB.LS were induced with 0, 0.4, 3.2, 12.8, and 19.2 μM Ni^{2+} for 4 hours.

To try and quantify the effect of using different concentrations of Ni^{2+} on the expression of LS, Western blot analysis (methods) was performed using anti-HA antibodies and PsaD-antibodies, which acted as a loading control and binded to the endogenous PSI subunit.

Expression was tested using 6803.nrsB.LS grown to an OD_{730} of 0.8, split into seven flasks, and induced for 4 hours with different concentrations of Ni^{2+} (Figure 3.11). The result confirmed that expression of LS is weak after induction with 0.4 and 0.8 μM Ni^{2+}, but is increased when 3.2, 6.4 and 12.8 μM Ni^{2+} is used for induction. Quantitative analysis using PsaD as the standard, confirmed that this was the case (Figure 3.11B). Expression peaks when 3.2 μM Ni^{2+} was used for induction and levels start to decrease at higher concentrations, possibly because of the inhibitory effect on growth of high nickel concentrations. To confirm this, another Western blot was run with the same samples (including an extra sample induced with 1.6 μM Ni^{2+} in duplicate (Figure 3.12). However, the quantitative analyses from both blots were not entirely in agreement (Figure 3.12). The first blot suggests that the expression peaked when 3.2 μM Ni^{2+} was used, however, in the second blot the expression peaked when 6.4 μM Ni^{2+}. This demonstrates the difficulty of using band quantification following Western blot analysis to accurately determine expression. Nevertheless, the study revealed that a range of between 3.2 and 6.4 μM Ni^{2+} is suitable to induce maximal levels of LS.

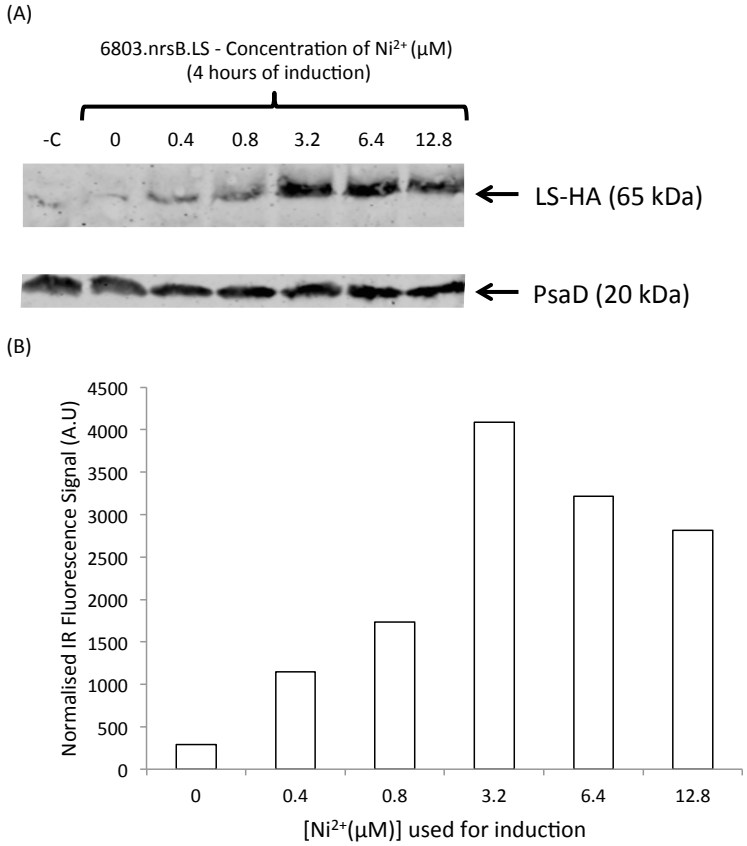


Figure 3.11 Western blot analysis comparing the effect different concentrations of Ni^{2+} has on the expression of LS under the control of P_{nrsB}.

(A) Western blot performed with whole cell extracts and anti-HA and anti-PsaD antibodies, using the Odyssey® Infrared Imaging System for detection. The samples were run on a 12% SDS-PAGE gel. The expected band size for LS is 65 kDa and 20 kDa for PsaD, which acted as a loading control. Protein size was determined using the PageRuler™ Prestained Protein ladder (Thermo Scientific). The strain, 6803.nrsB was used as negative control. (B) LS was quantified using the Odyssey® Infrared Imaging System. The LS IR fluorescence signal was normalised against the IR fluorescence signal of PsaD.

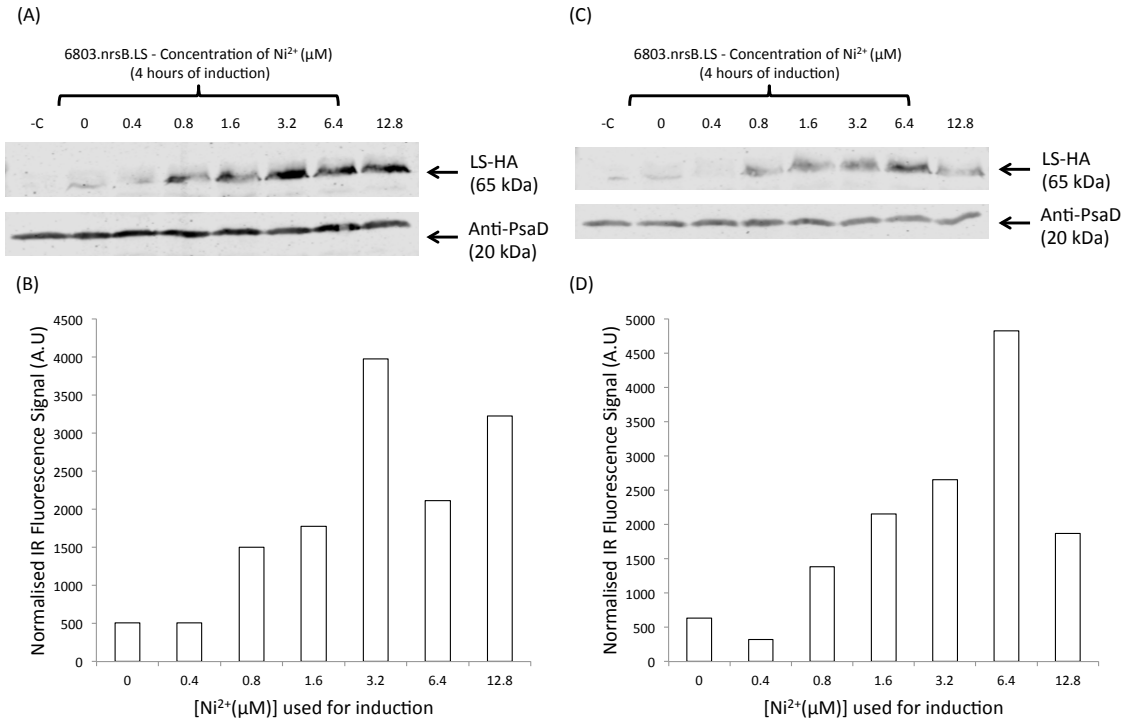


Figure 3.12 Western blot analysis, repeating the Western blot in Figure 3.11, comparing the effect different concentrations of Ni^{2+} has on the expression of LS under the control of P_{nrsB}.

(A) and (C) Western blot performed with whole cell extracts and anti-HA and anti-PsaD antibodies, using the Odyssey® Infrared Imaging System for detection. The samples were run on a 12% SDS-PAGE gel. The expected band size for LS is 65 kDa and 20 kDa for PsaD, which acted as a loading control. Protein size was determined using the PageRuler™ Prestained Protein ladder (Thermo Scientific). The strain, 6803.nrsB was used as a negative control. (B) and (D) LS was quantified using the Odyssey® Infrared Imaging System. The LS IR fluorescence signal was normalised against the IR fluorescence signal of PsaD.

3.2.2.3.2 Growth analysis on the effect of nickel on 6803.nrsB.LS and 6803.nrsB

The results of the toxicity tests suggest that a low concentration of limonene is toxic towards *Synechocystis* when added exogenously. If the same toxic effect is seen when limonene is produced within the cell, the growth of *Synechocystis* strains producing limonene will be affected. When preparing the cultures for Western blot analysis (section 3.2.2.3.1), after the addition of Ni^{2+}, the OD_{730} would decrease slightly in the 6803.nrsB.LS strains. To determine whether this effect was due to Ni^{2+} or the production of limonene, the growth of 6803.nrsB was compared to the growth of 6803.nrsB.LS before and after the addition of Ni^{2+} (methods) (Figure 3.13). 6803.nrsB without the addition of Ni^{2+} was also grown alongside as a control strain. The slower growth rate of 6803.nrsB in the absence of

Ni^{2+} was due to problems with the incubator, and was grown at around 26°C instead of 30°C.

The results indicate that the addition of nickel does have a negative effect on the growth of *Synechocystis*, as both strains are affected by the addition of Ni^{2+} whereas the growth of the control strain is unaffected. After the addition of 6.4 μM Ni^{2+}, photographs of the cultures were taken at 68.5 hours and confirmed the decrease in growth after the addition of Ni^{2+}. Clumps appeared in cultures where Ni^{2+} was added. Transition metal ions such as Ni^{2+} are required by enzymes to provide the catalytic activity. The *nrsBACD* and *nrsRS* operon present in *Synechocystis*, where the *nrsB* promoter originates from, is responsible for maintaining Ni^{2+} homeostasis (García-Domínguez et al., 2000; López-Maury et al., 2002). However, Ni^{2+} can inhibit photosynthesis and respiration at high concentrations (Kleiner, 1978; Rai et al., 1994).

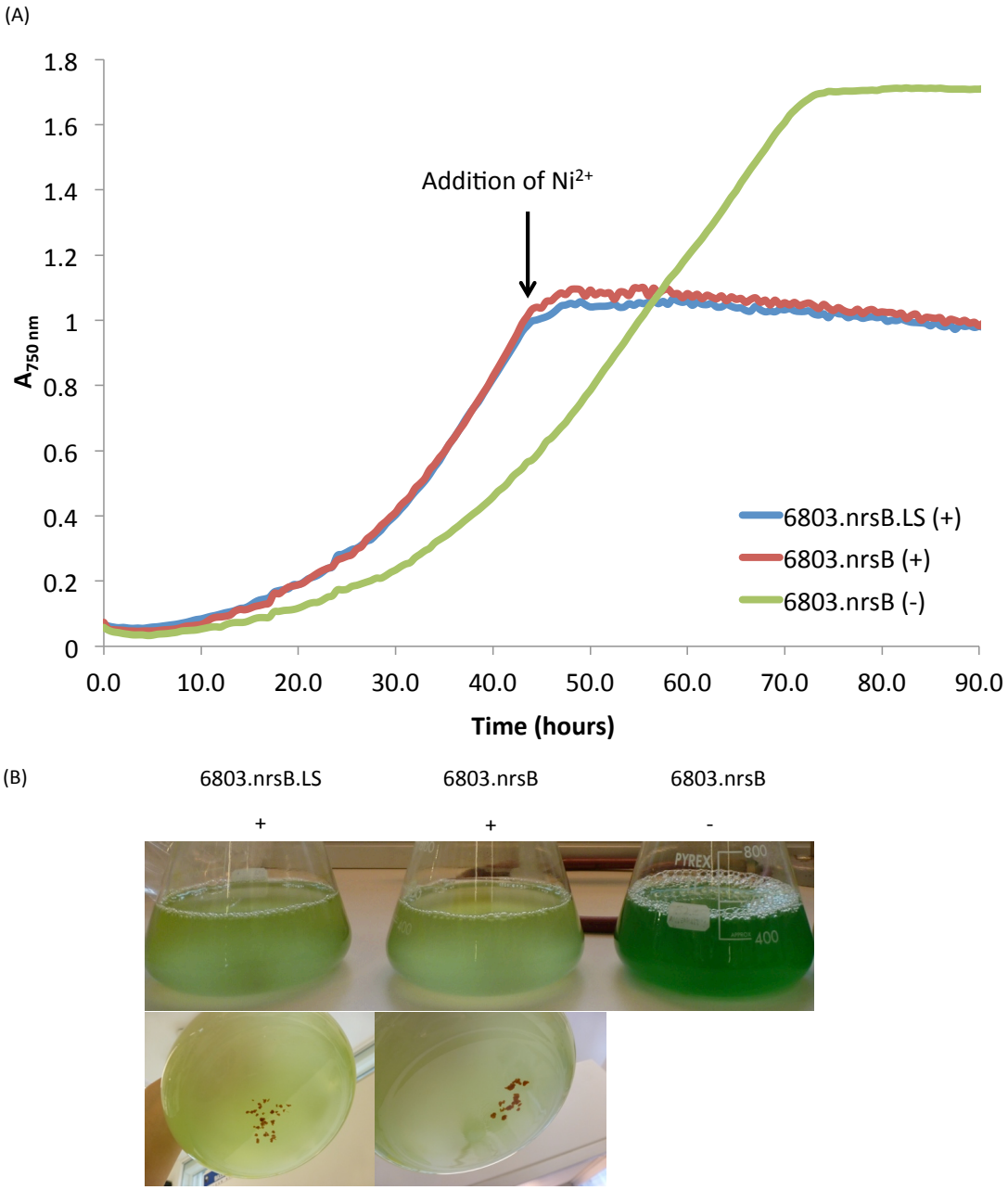


Figure 3.13 Comparison of cell growth between 6803.nrsB.LS and 6803.nrsB after induction with Ni^{2+}.

(A) The growth, measured by the A_{750}, of strains 6803.nrsB.LS and 6803.nrsB.LS and 6803.nrsB was monitored in the Algem photobioreactor (200 µmol/m^2/s white light, at 30°C and 120 rpm shaking) before and after induction with 6.4 µM Ni^{2+} after 43 hours (+). The strain 6803.nrsB was also grown without the addition of Ni^{2+} (−). (B) Photographs were taken after 68.5 hours.

3.2.2.4 Creation of transgenic strains of *Synechocystis* containing the limonene synthase gene under the control of the P_{psbAII} promoter

The induction of LS resulted in the accumulation of the protein (Figure 3.10) but did not result in the death of the culture, seen when limonene was added to the medium (Figure 3.3), suggesting that any LS activity produced in the transformant cells is below toxic levels. It was therefore decided to further increase the production of LS in the cell, by putting the gene under the control of the stronger promoter, P_{psbAII}.

As before, *ls* was digested from ls_pMA-RQ using NdeI and BamHI and cloned into the NdeI and BamHI site of expression vector pLAH.AII (see section 2.3 for details) to create pLAH.AII.LS (Figure 3.14).

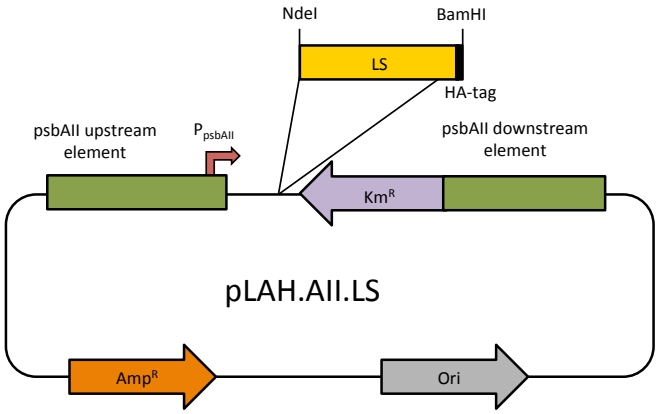


Figure 3.14 Diagram of the expression plasmid, pLAH.AII.LS

The *ls* gene digested from ls_pMA-RQ was cloned into pLAH.AII to create pLAH.AII.LS, to enable the integration of *ls* under the control of P_{psbAII} at the *psbAII* site.

The plasmid was checked both by restriction digest and sequencing, and then used to transform WT *Synechocystis*. After selection with kanamycin, eight colonies were picked for four rounds of selection on kanamycin-containing media. PCR was performed on putative 6803.AII.LS transformants (Figure 3.15). All eight putative transformants tested showed the integration of *ls* under P_{psbAII} at the *psbAII* site, and were judged to be homoplasmic.

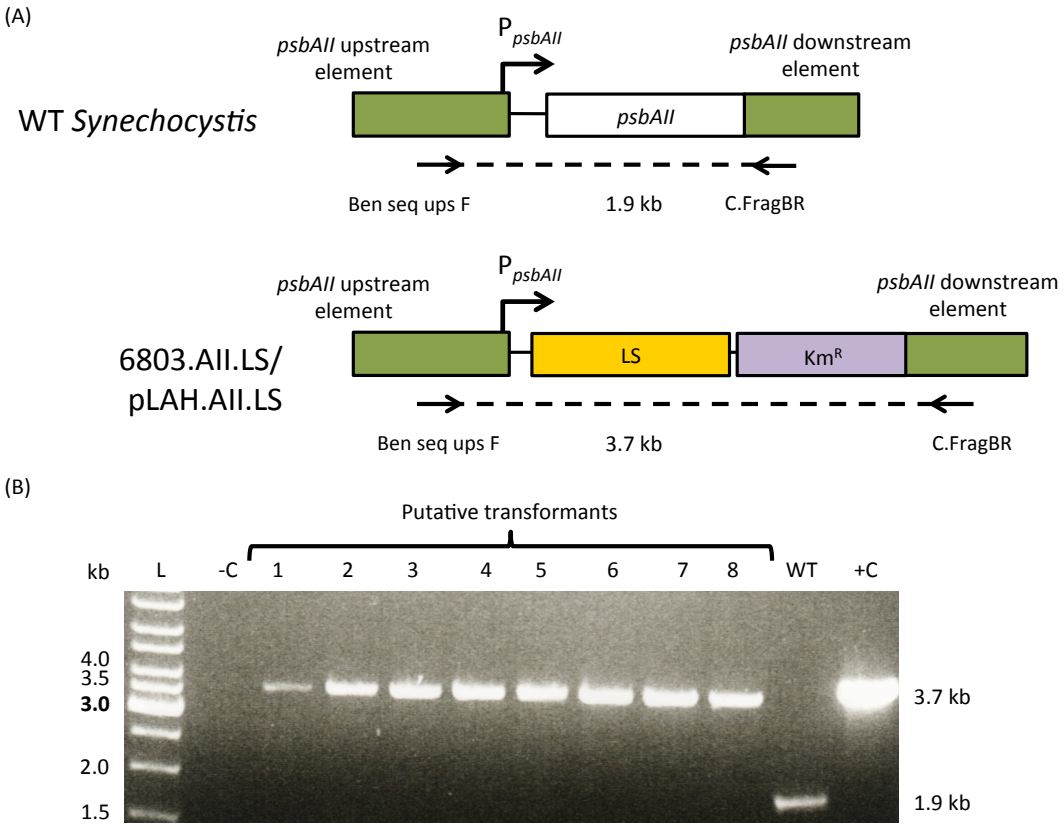


Figure 3.15 PCR screening for successful transformants of *Synechocystis* containing *ls* under the control of P_{psbAII}.

(A) Schematic showing the binding sites and expected fragments sizes for WT *Synechocystis*, pLAH.AII.LS expression plasmid and successful 6803.AII.LS transformants. (B) Colony PCR screening of putative transformants of 6803.AII.LS. WT *Synechocystis* (WT), no DNA (-C) and pLAH.AII.LS (+C) were used as controls. The single band at 3.7 kb confirmed homoplasmic strains were produced.

3.2.2.5 Expression of limonene synthase under the control of the P_{psbAII}

Western blot analysis was performed with anti-HA antibodies on the crude extract of three transformants of 6803.AII.LS using the Odyssey® Infrared Imaging System to detect the expression of LS (Figure 3.16). Also included were samples of 6803.nrsB.LS that had been induced for 4h with different concentrations of Ni^{2+}. The expression of LS was seen in all three 6803.AII.LS transformants, and all 6803.nrsB.LS samples, that were induced with Ni^{2+}. From the Western blot, the expression of LS under the P_{nrsB}, at all concentrations, is seen to be significantly weaker than the expression of LS under the control of P_{psbAII}.

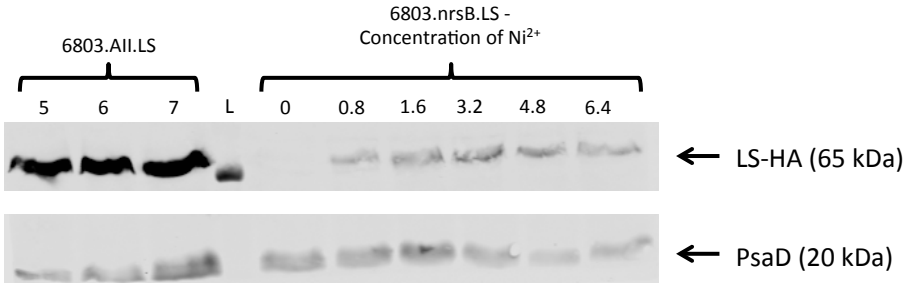


Figure 3.16 Western blot analysis confirming the expression of HA-tagged LS in 6803.AII.LS.

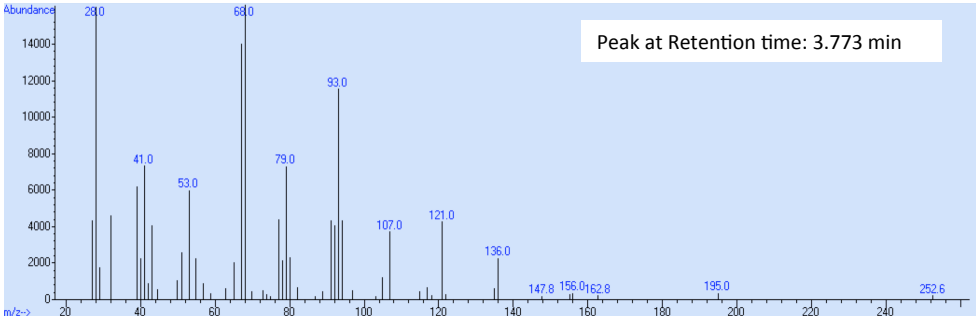
Western blot analysis was performed with whole cell extracts with anti-HA and anti-PsaD antibodies using the Odyssey® Infrared Imaging System for detection. Samples were run on a 12% SDS-PAGE gel. Protein size was determined using the PageRuler™ Pre-stained Protein ladder (L) (Thermo Scientific™). The expression of three transformants of 6803.AII.LS was compared against the expression of 6803.nrsB.LS induced with different concentrations of Ni^{2+}. The expected band size for LS was 65 kDa and 20 kDa for PsaD, which acted as a loading control.

3.2.2.6 Limonene extraction and GC-MS analysis

To test whether strains of 6803.AII.LS were able to successfully produce limonene, two methods were used to recover any limonene produced from the strains. One method used ethyl acetate to extract limonene directly from cells and the other method used a two-phase extraction with a 10% (v/v) overlay of dodecane over a growing culture of 6803.AII.LS. Limonene standards of S-limonene prepared in ethyl acetate and dodecane were also prepared.

GC Analysis was performed on the two extracts. Limonene was readily detected when using the standard (50 mg/L S-limonene in ethyl acetate), as a single peak at 3.773 minutes, and gave peak fragments at 136, 121, 93 and 68 m/z as found in the library standard of limonene (Figure 3.17). However, limonene could not be detected in any of the 6803.AII.LS samples (Figure 3.18).

(A)



(B)

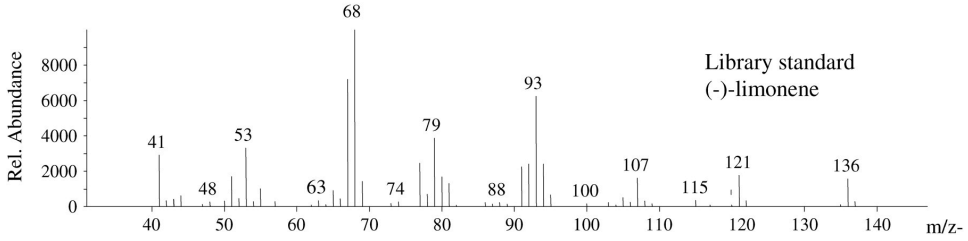


Figure 3.17 GC-MS analysis of 5 mg/L limonene in dodecane

(A) The mass spectra of the compound found at 3.773 retention time has the same characteristic peaks of at 68, 93, 121 and 136 m/z as the (B) library standard mass spectra of (-)-limonene from (Byun-McKay et al., 2006).

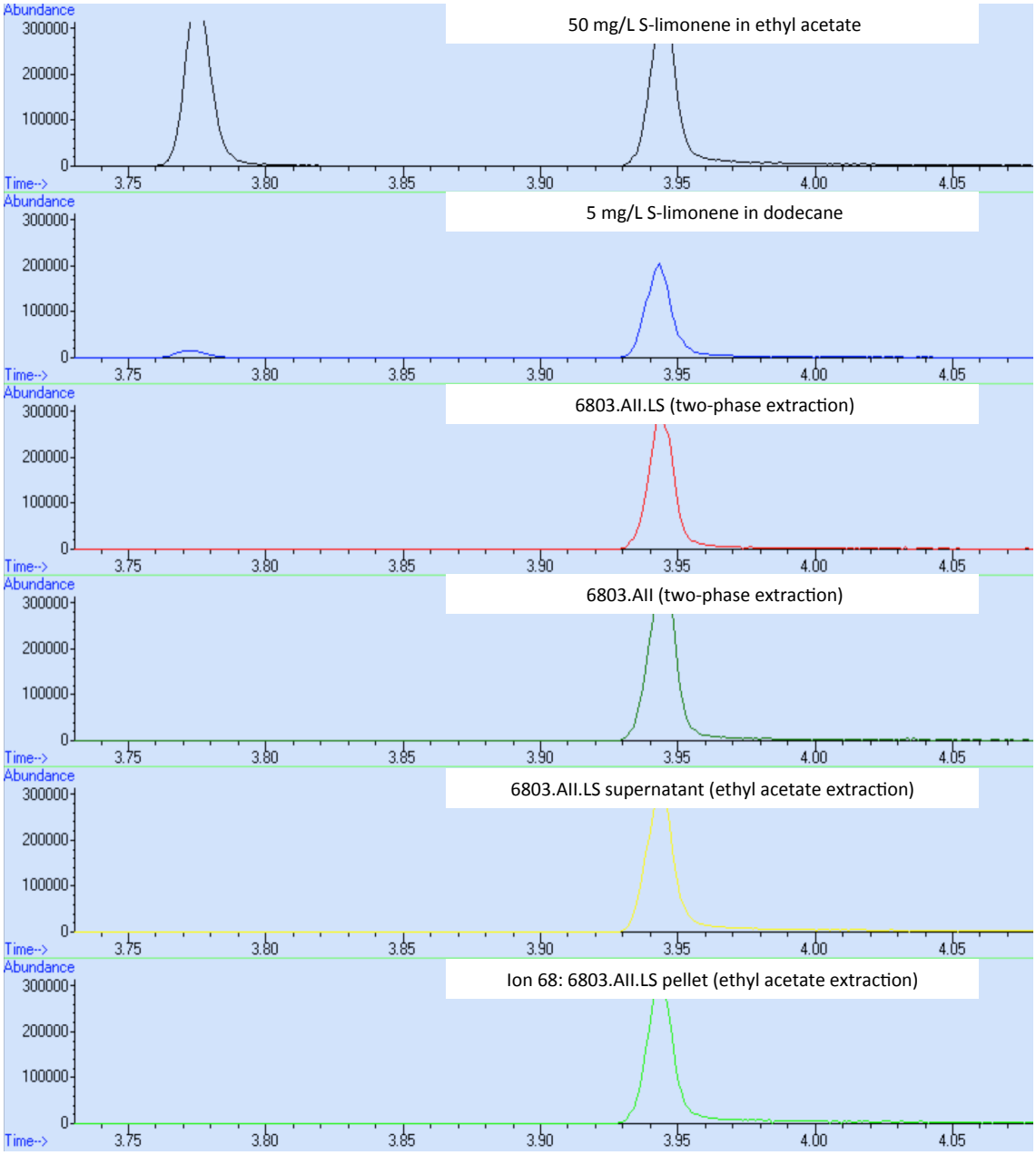


Figure 3.18 GC-MS analyses of dodecane overlay-extracted from 6803.AII.LS and ethyl acetate extraction of 6803.AII.LS.

MS chromatogram (ion overlay at 68 m/z) of the samples extracted from 6803.AII.LS by using dodecane and ethyl acetate. The peak at 3.773 minutes indicates limonene presence and the peak at ~3.94 minutes is the internal standard, 1-octanol. No limonene is detected in any of the 6803.AII.LS samples.

3.2.2.7 Glass beads method for extraction

One possible explanation for the lack of limonene detection may be that the *ls* gene designed did not result in a functional gene product. To assess whether this was the case, an enzyme assay testing the recombinant LS was attempted. In order to perform an enzyme assay using crude cell extracts, it was first necessary to find an alternative method

of breaking open the cells without using heat, which can damage the enzyme. To extract the crude protein extract, the samples are usually boiled to break open the cells when using the 'Heat' method (see 2.5.1 and 2.5.3 for details). Glass beads could also be used as an alternative method of breaking the cells open (see 2.5.2 and 2.5.3 for details). When using this method, the soluble proteins can be separated from the membrane proteins, present in the lower fraction.

A Western blot analysis was performed on extracts from single transformant of 6803.AII.LS (grown to an OD_{730}: 0.603) and 6803.nrsB.LS (grown to an OD_{730}: 0.481), after they were prepared using the two methods (Figure 3.19). As seen in the figure, glass beads are an effective method of breaking the cells open, with LS detected mostly in the supernatant (although some LS is detected in the pellet where the membrane proteins are expected to be present). When *ls* is expressed under P_{psbAII}, the amount of LS in the supernatant was similar to the total amount of LS recovered following the heating method. The results also confirm that the expression of LS under the control of P_{psbAII} is much greater than under the control of P_{nrsB} following a 4 hour induction with 6.4 μM Ni^{2+}.

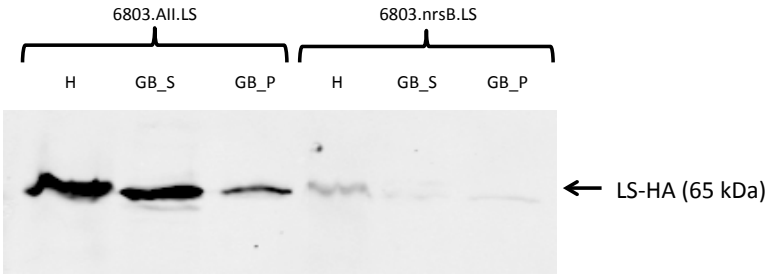


Figure 3.19 Western blot analysis confirming the detection of LS using Glass beads as a method of cell disruption on strains 6803.AII.LS and 6803.nrsB.LS.

Western blot analysis was performed on extracts of 6803.AII.LS and 6803.nrsB cells broken with glass beads using the Odyssey® Infrared Imaging System for detection. Samples were run on a 12% SDS-PAGE gel. The upper layer of extract when using glass beads was referred to as the supernatant (GB_S) and the lower layer the pellet (GB_P). Protein size was determined using the PageRuler™ Pre-stained Protein ladder (Thermo Scientific). The detection of LS using heat (H) as the method of cell disruption acted as the positive control.

3.2.2.8 Creation of transgenic strains of *Synechocystis* containing limonene synthase under the control of the P_{cpc560}

As presented in detail in Chapter 4, the promoter P_{cpc560} is stronger than P_{psbAII}, therefore the expression of *ls* under the P_{cpc560} was attempted in *Synechocystis* to further increase expression and thereby improve the chances of detecting LS activity.

The *ls* gene was again digested from ls_pMA-RQ using NdeI and BamHI, and cloned into the NdeI and BamHI restriction sites of pLAH.cpc (see 2.3 for more details) to create pLAH.cpc.LS (Figure 3.20). The creation of pLAH.cpc.LS was confirmed by restriction digest and sequencing analysis.

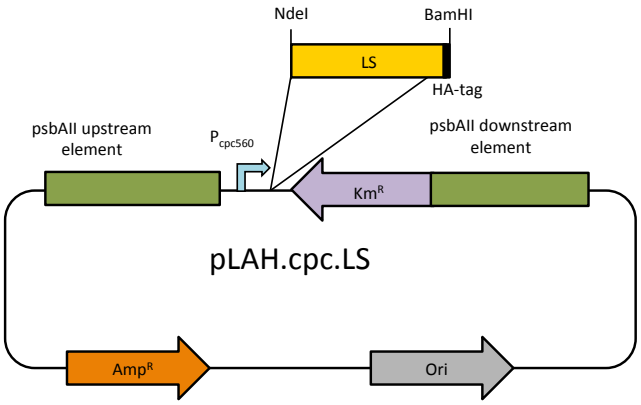


Figure 3.20 Diagram of the expression plasmid, pLAH.cpc.LS

The *ls* gene digested from ls_pMA-RQ was cloned into the pLAH.cpc to create pLAH.cpc.LS, to enable the integration of *ls* under the control of P_{cpc560} at the *psbAII* site.

WT *Synechocystis* was transformed with pLAH.cpc.LS. After selection with kanamycin, colonies appeared after 10 days and eight colonies were picked for another round of selection on kanamycin-containing media. All putative transformants were screened by PCR and confirmed the insertion of *ls* under the control of P_{cpc560} at the *psbAII* site, to produce 6803.cpc.LS (Figure 3.21). All transformants were homoplasmic as all copies of *psbAII* have been replaced with *ls*.

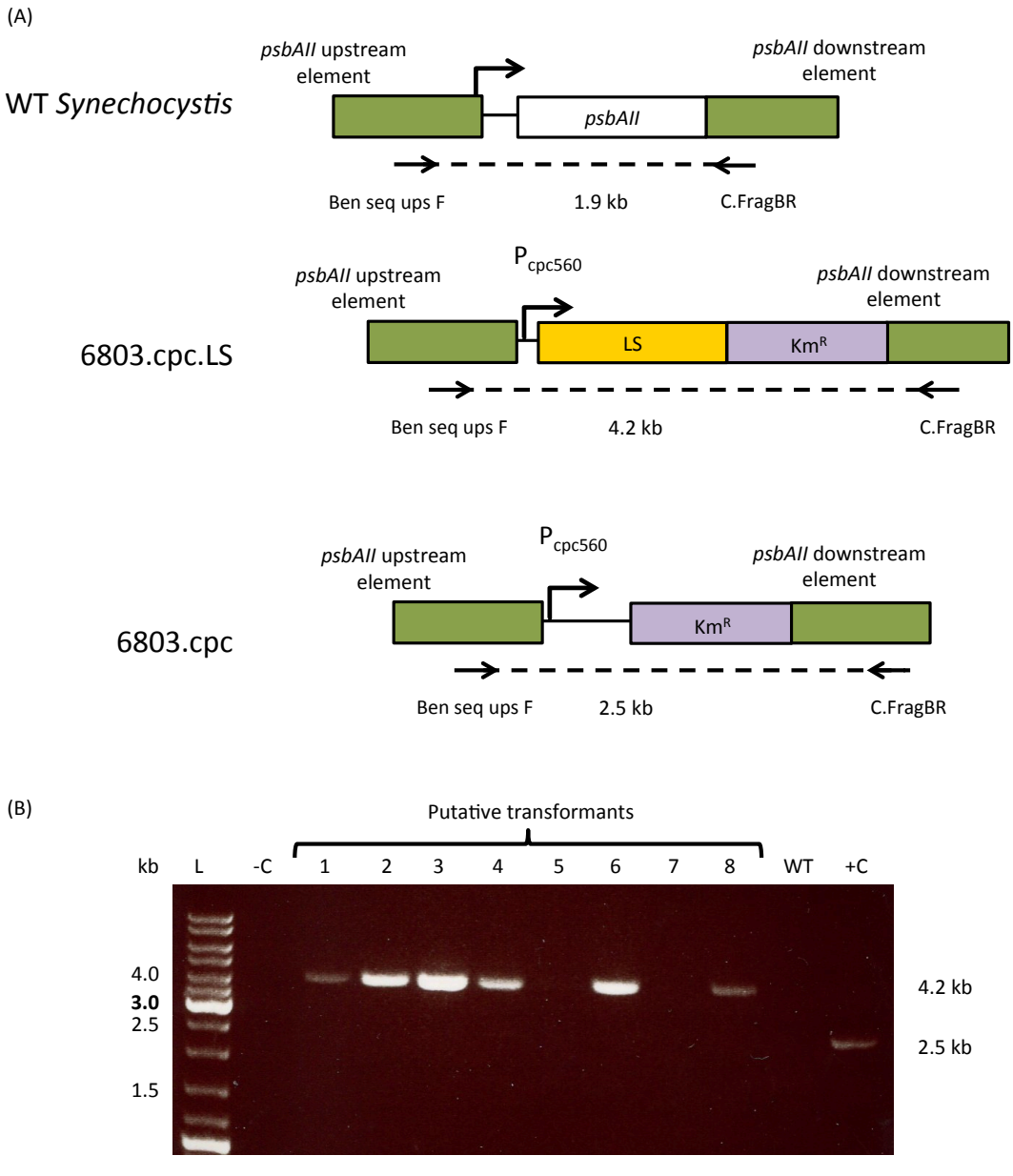


Figure 3.21 PCR screening for successful transformants of *Synechocystis* containing *ls* under the control of $P_{cpcp560}$.

(A) Schematic showing the binding sites and expected fragments sizes for WT *Synechocystis*, 6803.cpc and successful 6803.cpc.LS transformants. (B) PCR screening of putative transformants of 6803.cpc.LS. WT *Synechocystis* (WT), no DNA (–C) and 6803.cpc (+C) were used as controls. The single band at 4.2 kb confirmed homoplasmic strains of 6803.cpc.LS.

3.2.2.9 Expression of limonene synthase under the control of the P_{cpc560}

The expression of LS under the P_{cpc560} promoter in three strains of 6803.cpc.LS was confirmed by Western blot analysis, using anti-HA antibodies and the Odyssey® Infrared Imaging System for detection (Figure 3.22). Expression of LS under the control of the P_{cpc560} appeared to be greater than expression under the control of the P_{psbAII} (Figure 3.22).

Although the level of LS is hard to quantify because the blot appears to have moved during blotting or exposure, the amount of LS does appear to be higher when the gene is under the P_{cpc560} promoter.

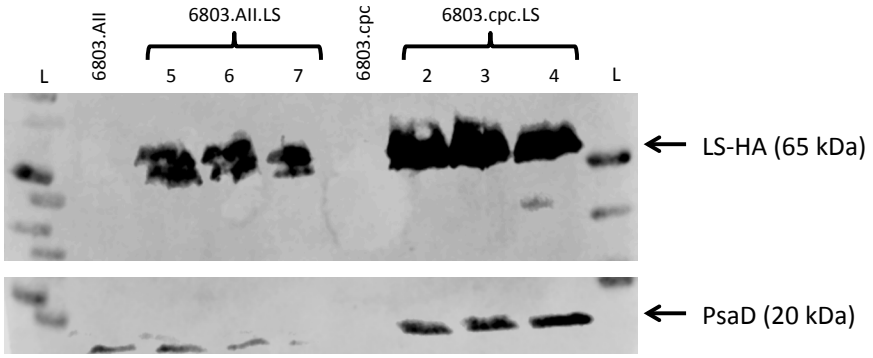


Figure 3.22 Western blot analysis confirming the expression of LS in 6803.cpc.LS and comparing the expression with 6803.AII.LS

Western blot analysis was performed with whole cell extracts with anti-HA and anti-PsaD antibodies using the Odyssey® Infrared Imaging System for detection. Samples were run on a 12% SDS-PAGE gel. Protein size was determined using the PageRuler™ Pre-stained Protein ladder (L) (Thermo Scientific). The expression of three transformants of 6803.cpc.LS was compared against three transformants of 6803.AII.LS. The expected band size for LS is 65 kDa and 20 kDa for PsaD, which acted as a loading control. 6803.AII and 6803.cpc were loaded as negative controls.

3.2.2.9.1 Growth analysis comparing 6803.cpc.LS and 6803.cpc

As the expression of LS was greater when under the control of the P_{cpc560} promoter, the growth of 6803.cpc and 6803.cpc.LS was compared to see whether the greater expression of LS and the production of limonene had an effect on growth. A growth curve was performed in an Algem photobioreactor and the growth rates of 6803.cpc.LS with 6803.cpc (Figure 3.23) were also calculated. Both strains were grown in the same conditions in duplicate. Although the 6803.cpc.LS strain showed a longer lag phase in this particular experiment, both 6803.cpc and 6803.cpc.LS had an equivalent growth rate of 0.01 h^{-1} and both reached a similar OD at stationary phase. The expression of LS therefore has no effect on growth of 6803.cpc.LS. If limonene is being produced then it is not being produced to a high enough concentration to have a toxic effect on the cell.

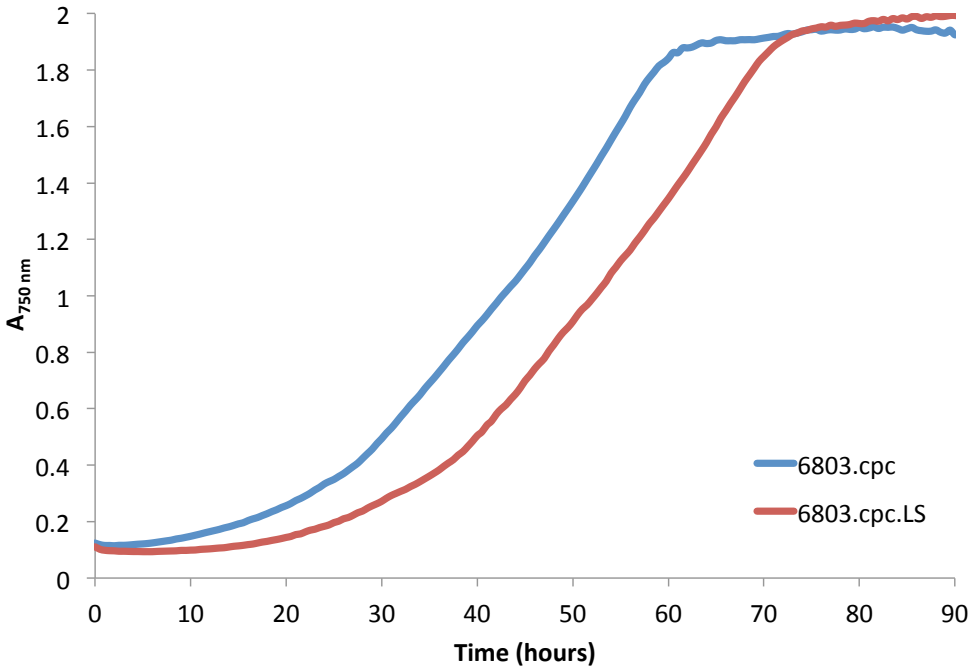


Figure 3.23 Comparison of cell growth between 6803.cpc and 6803.cpc.LS

The growth, measured by the A_{750}, of strains 6803.cpc and 6803.cpc.LS were monitored using the Algem photobioreactor (200 µmol/m^2/s white light, at 30°C and 120 rpm shaking).

3.2.2.10 Enzyme assay and DPPH assay for limonene production

As the production of limonene was not detected in 6803.AII by GC-MS analysis, an enzyme assay on LS expressed from 6803.cpc.LS was performed. To try and detect the production of limonene without using GC-MS, the 2,2-Diphenyl-1-picrylhydrazyl (DPPH) assay was used to detect limonene production in *S. cerevisiae* with some modifications (Behrendorff et al., 2013). The assay uses DPPH, a stable radical that produces a strong absorbance at 517 nm, and at 510 nm when DPPH is dissolved in dodecane. When DPPH loses a radical in exchange for a proton from limonene, the absorbance at 510 nm decreases proportionately, and DPPH in dodecane goes from purple to yellow (Figure 3.24) (Behrendorff et al., 2013). As the DPPH assay is performed in dodecane, a dodecane overlay was used in the enzyme assay to trap any limonene produced by 6803.cpc.LS.

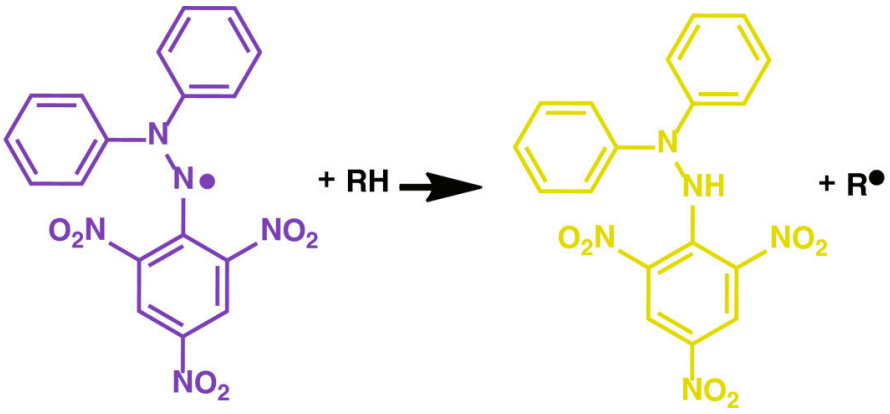


Figure 3.24 DPPH Reaction

DPPH exhibits strong absorbance at 517nm (Purple) but when it loses its radical for a proton, its absorbance decreases proportionally and becomes yellow. Figure from (Behrendorff et al., 2013)

3.2.2.10.1 Establish the DPPH assay

The DPPH assay was capable of detecting limonene at 200 µM, so initial tests were performed to try and emulate the assay (Behrendorff et al., 2013). Using the FLUOstar Omega plate reader the maximum absorbance of DPPH dissolved in dodecane was seen as a single peak at 510 nm and there was a slight decrease over time. To establish the lowest concentration of limonene that can be detected by the plate reader, the difference between the $A_{510\ nm}$/min of DPPH in dodecane was compared to different concentrations of limonene in DPPH (Figure 3.25). A threshold was employed using an unpaired Student's t-test, when $p<0.01(n=4)$, between the DPPH in dodecane and limonene in DPPH (Behrendorff et al., 2013). Using this threshold, limonene was significantly detected at 0.8 mM; making the assay in our hands less sensitive as it is four-times higher than the detection level seen by Behrendorff et al (Behrendorff et al., 2013). This could be due to oxidation of the limonene standards, as this decreases the proportion of limonene available to react with DPPH.

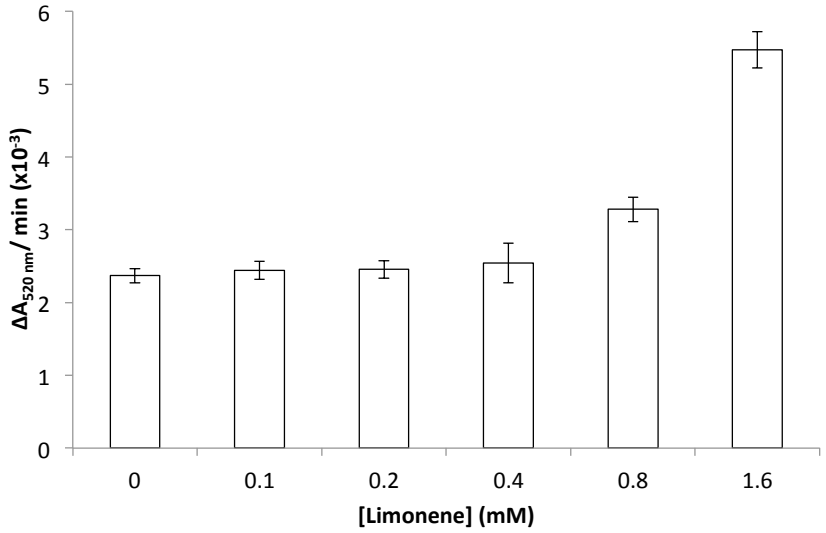


Figure 3.25 Detection of limonene after incubation with DPPH.

The reaction rates between 0.1 mM DPPH and varying concentrations of limonene were calculated between 10-180 min (n=4 for each concentration, mean ± 1 SD). Reaction rates for DPPH limonene standards were compared to the relevant control using an unpaired Student's *t*-test.

3.2.2.10.2 Enzyme assay with 6803.cpc.LS using DPPH assay for detection

The enzyme assay was designed to be performed using crude protein extract from the supernatant of 6803.cpc.LS with cells broken using the glass bead method. Western blot analysis confirmed the presence of LS in the supernatant of 6803.cpc.LS, to be used in the enzyme assay (Figure 3.26). GPP, the substrate of LS, would then be added to the supernatant and a layer of dodecane would then trap any limonene produced. The dodecane would then be used in the DPPH assay to see if limonene was produced. The DPPH assay was also performed with limonene standards of different concentrations in 0.1 mM DPPH.

The results of the DPPH assay suggests that the difference in the $\Delta A_{510\ nm}$/min for the control strain, 6803.cpc, is actually higher than the 6803.cpc.LS strain when 1 mM and 1.5 mM GPP are added to the cells (Figure 3.27). This suggests that limonene is not produced by 6803.cpc.LS strains to a detectable level using this assay. As duplicates of the enzyme assay were not performed the $\Delta A_{520\ nm}$/min of 6803.cpc and 6803.cpc.LS (+0, 1, 1.5 mM GPP) were compared to the standards run alongside the enzyme assay (Figure 3.27). The $\Delta A_{520\ nm}$/min values, from the enzyme assay were all below the values seen when 0.8 mM limonene is present, the threshold for limonene detection by the plate reader. This

confirms that the $\Delta A_{520\ nm}$/min values seen in the enzyme assay were not caused by the production of limonene.

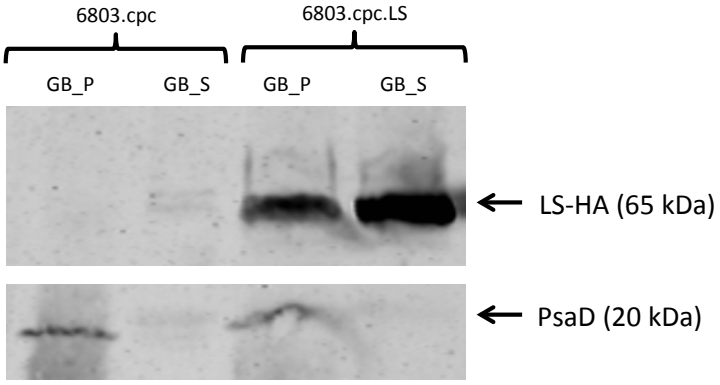


Figure 3.26 Western blot analysis confirming the expressing of LS in the supernatant used in the enzyme assay.

Western blot analysis was performed with whole cell extracts with anti-HA and anti-PsaD antibodies using the Odyssey® Infrared Imaging System for detection. Samples were run on a 12% SDS-PAGE gel. Protein size was determined using the PageRuler™ Pre-stained Protein ladder (Thermo Scientific). LS was detected in the supernatant (GB_S) and pellet (GB_P) of 6803.cpc.LS. The expected band size for LS is 65 kDa and 20 kDa for PsaD, which acted as a loading control.

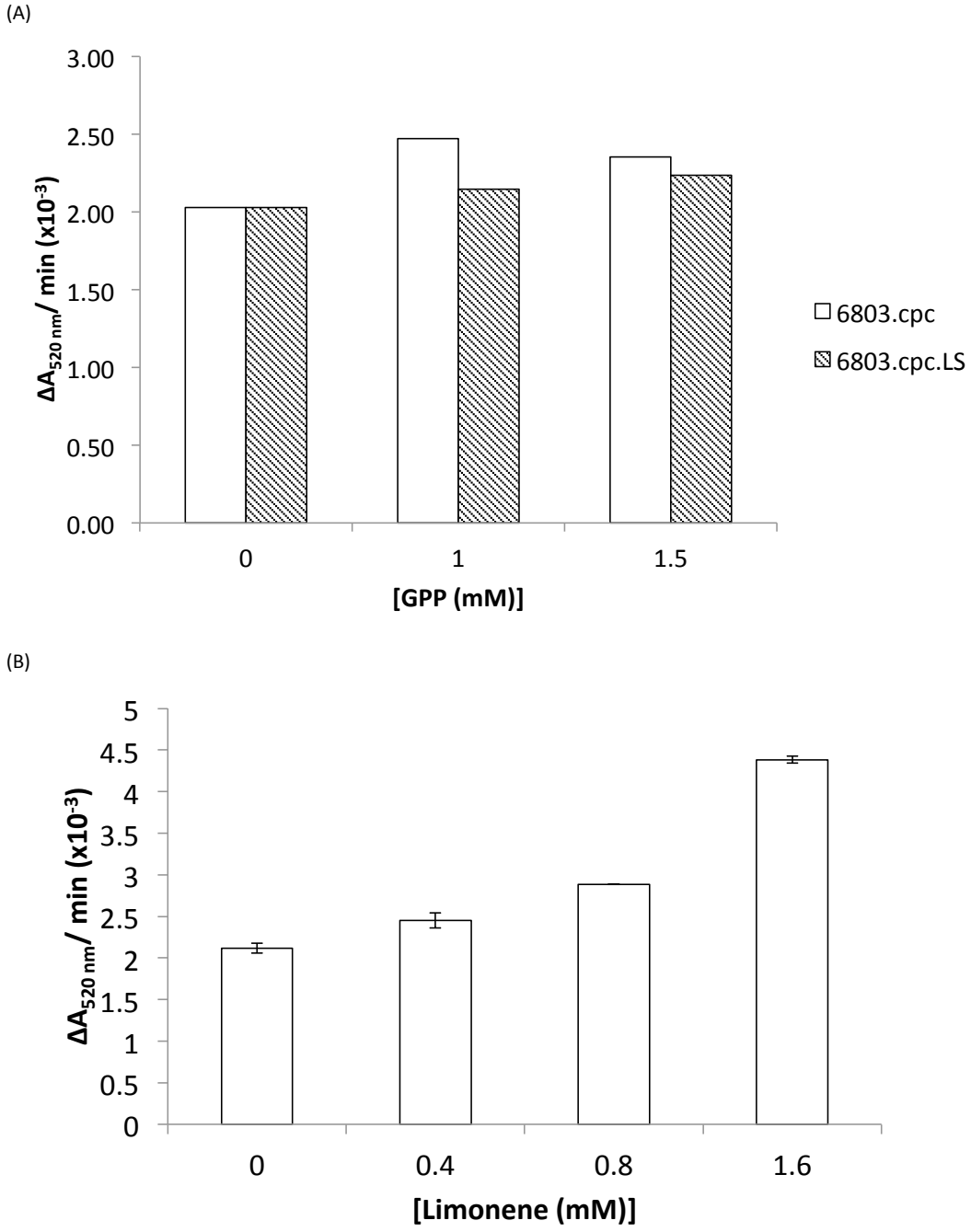


Figure 3.27 Reaction rates of the dodecane layer from the enzyme assay and limonene standards after incubation with DPPH.

(A) Reaction rates of the dodecane layer from the enzyme assay performed on the supernatant of 6803.cpc and 6803.cpc.LS with 0, 1 and 1.5 mM GPP, after incubation with 0.1 mM DPPH. (B) Reaction rates between 0.1 mM DPPH and varying concentrations of limonene were calculated between 10-180 min (n=3 for each concentration (except 1.6 mM n=2), mean ± 1 SD). Reaction rates for DPPH limonene standards were compared to the relevant control using an unpaired Student's *t*-test.

3.2.3 Increasing limonene production: deletion of *sds*

One possible explanation for the lack of detectable limonene in the transformants despite the presence of the LS enzyme is that the pool of available GPP substrate is low in the cell or inaccessible to the LS. An attempt was therefore made to increase the pool of GPP for LS by deleting a gene downstream of GPP in the pathway for the terpenoid backbone synthesis. The KEGG pathway of *Synechocystis* – 'terpenoid backbone synthesis' – was reviewed when finding possible genes downstream of GPP that could be deleted. One identified gene (*sds*: gene *slr0611*) was that encoding solanesyl diphosphate synthase (SDS), which produces solanesyl diphosphate, a precursor of the side chains in ubiquinone-9 (UQ-9)/plastoquinone-9 (PQ-9) (Figure 3.28) (Kanehisa and Goto, 2000; Okada et al., 1997; Sadre et al., 2012). According to the on-line database CyanoBase the deletion of *slr0611* has been achieved by inserting a kanamycin resistance cassette at the SmaI site within the gene, and this resulted in no phenotypic change in (Nakamura et al., 1999). Based on this evidence, *sds* was chosen for deletion in *Synechocystis*, in an attempt to increase the pool of GPP for LS.

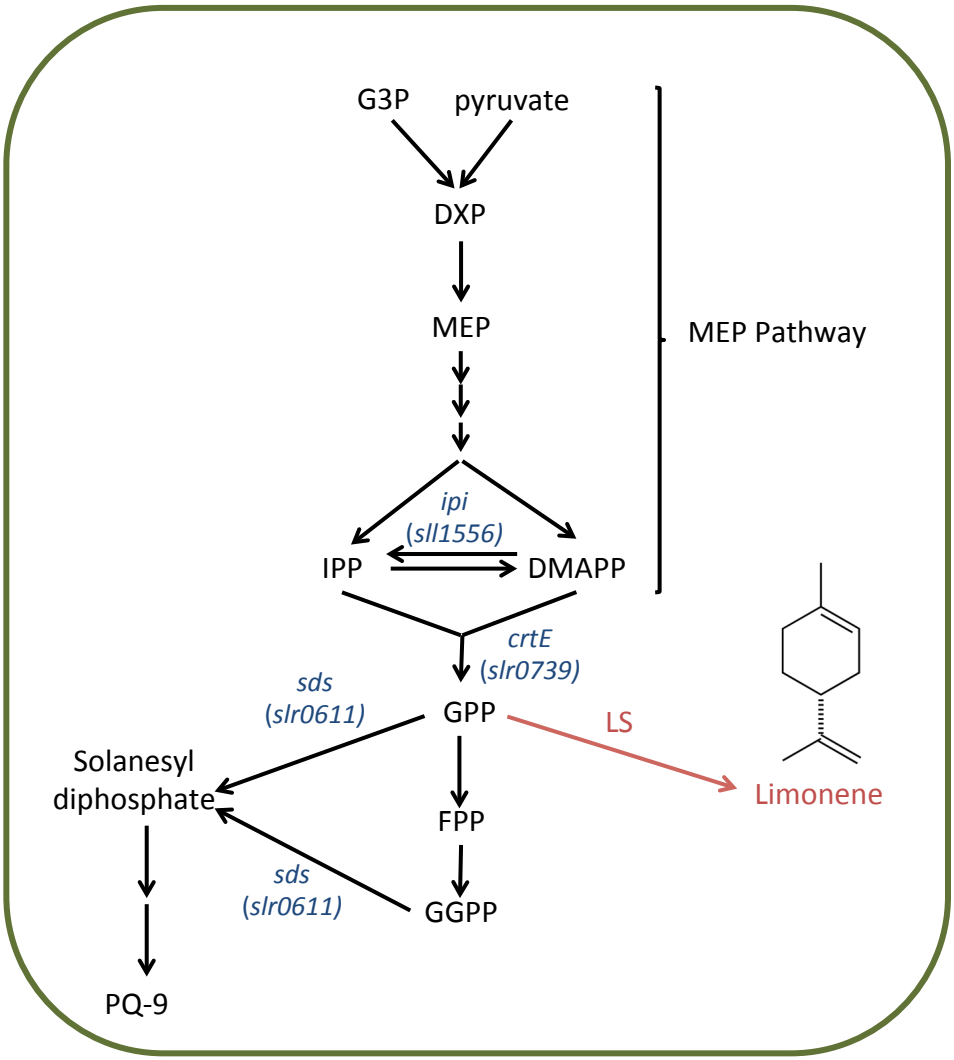


Figure 3.28 MEP pathway and pathway downstream of GPP for the production of PQ-3.

Schematic showing the methyerythritol 4-phosphate (MEP) pathway PQ-3 pathway in *Synechocystis.* Native genes in *Synechocystis* in blue and non-native pathway in red. Abbreviations used: G3P = glyceraldehyde 3-phosphate, DXP = 1-deoxy-d-xylulose 5-phosphate, MEP = methylerythritol 4-phosphate, IPP = isopentenyl diphosphate, DMAPP = dimethylallyl diphosphate, GPP = geranyl diphosphate, LS = limonene synthase, PQ-9 = Plastoquinone-9.

3.2.3.1 Creation of an expression plasmid to knockout *sds* in *Synechocystis*

To knockout *sds*, the expression plasmid, pJET.sds::Km was created (Figure 3.29).

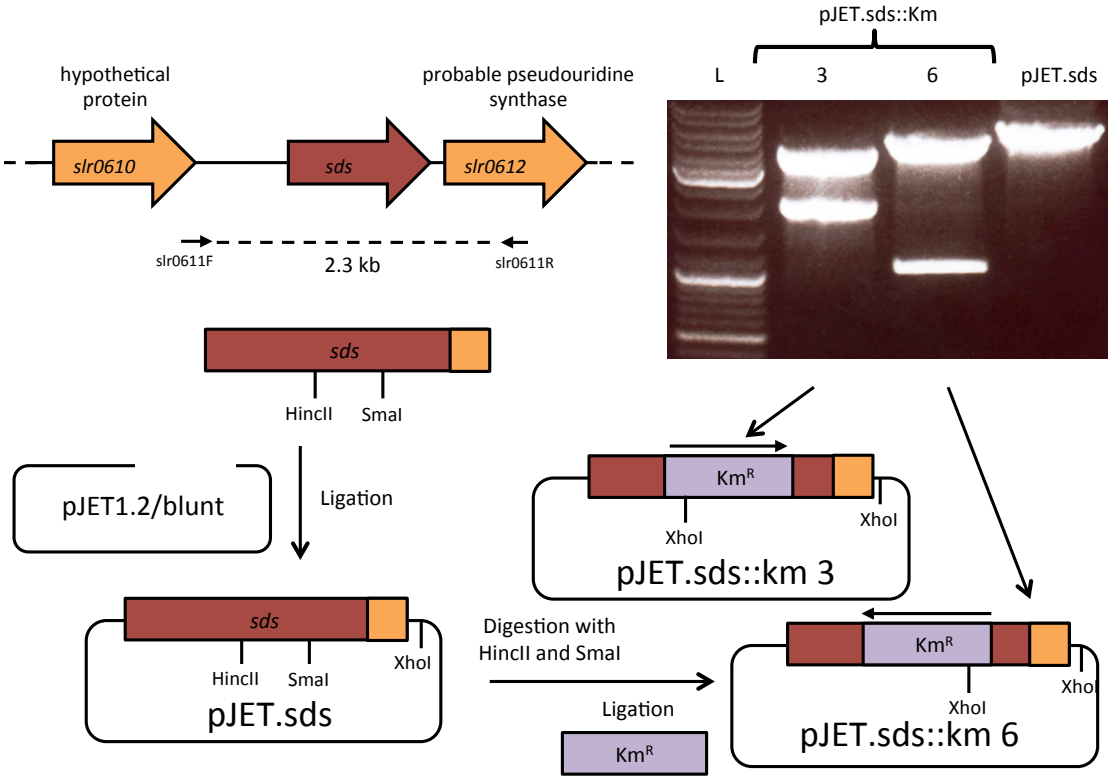


Figure 3.29 Diagram showing to creation of pJET.sds::km and confirmation of its creation by restriction digest.

Schematic showing the creation of the expression plasmid, pJET.sds::km, with the kanamycin resistance cassette in two different orientation.

Primers, slr0611F and slr0611R were designed and successfully used to amplify *sds* and approximately 500-800 bp upstream and downstream of *sds* from the genome of *Synechocystis* via PCR. The PCR product was cloned into the blunt cloning site of pJET1.2/blunt (see appendix) to produce pJET.sds. A restriction digest with HincII and SmaI and sequencing results with primers, pJETR and pJETF, confirmed the insertion of *sds* and creation of pJET.sds.

A kanamycin resistance cassette, Km^R, was digested from pUC4K (see appendix), using HincII, and cloned into pJET.sds at the SmaI and HincII blunt restriction sites to successfully create pJET.sds::Km. A restriction digest with XhoI confirmed the construction of pJET.sds::Km. Two variations of pJET.sds::Km were created, pJET.sds::Km3 and pJET.sds::Km6, as Km^R was inserted in the two possible orientations at the blunt-ended restriction sites (Figure 3.29). Sequencing with FO.kmF and FO.kmR, confirmed the insertion of the Km^R into pJET.sds, in the two different orientations.

3.2.3.2 Attempts to knockout *sds* in *Synechocystis*

WT *Synechocystis* was transformed with pJET.sds::Km3 and after selection on kanamycin-containing media, four colonies were picked after 14 days. After four rounds of selection on medium supplemented with kanamycin a PCR was performed using primers, slr0611F and slr0611R. All four transformants appeared to have the Km^R gene integrated at the *sds* locus, however all transformants still appeared to have copies of the wild-type *sds*. After four more rounds of selection on kanamycin, the strain remained heteroplasmic, and copies of WT *sds* remained in the genome. The putative transformants were then streaked onto BG-11 medium with and without the antibiotic, to see whether the strain would revert back to WT in the absence of selection for Km^R. After 10 rounds of selection on kanamycin, the *sds* gene still remained, indicating that the gene is essential for *Synechocystis* to grow (Figure 3.30B). However, after removing kanamycin from the medium, the Km^R gene also remained, suggesting that the integration was relatively stable.

WT *Synechocystis* was also transformed with pJET.sds::Km6 to see whether the orientation of the Km^R gene had an effect on the genes flanking *sds*, as *sds* may not be essential but one of the genes flanking *sds* may be essential instead. After selection with kanamycin, eight putative transformants were picked. After eight rounds of selection on kanamycin, PCR was performed as before. Similarly to the previous PCR, these new transformants also retained copies of both Km^R and *sds* (Figure 3.30C). When restreaked onto BG-11 medium, after 10 rounds without selection, the Km^R gene was eventually lost (results not shown).

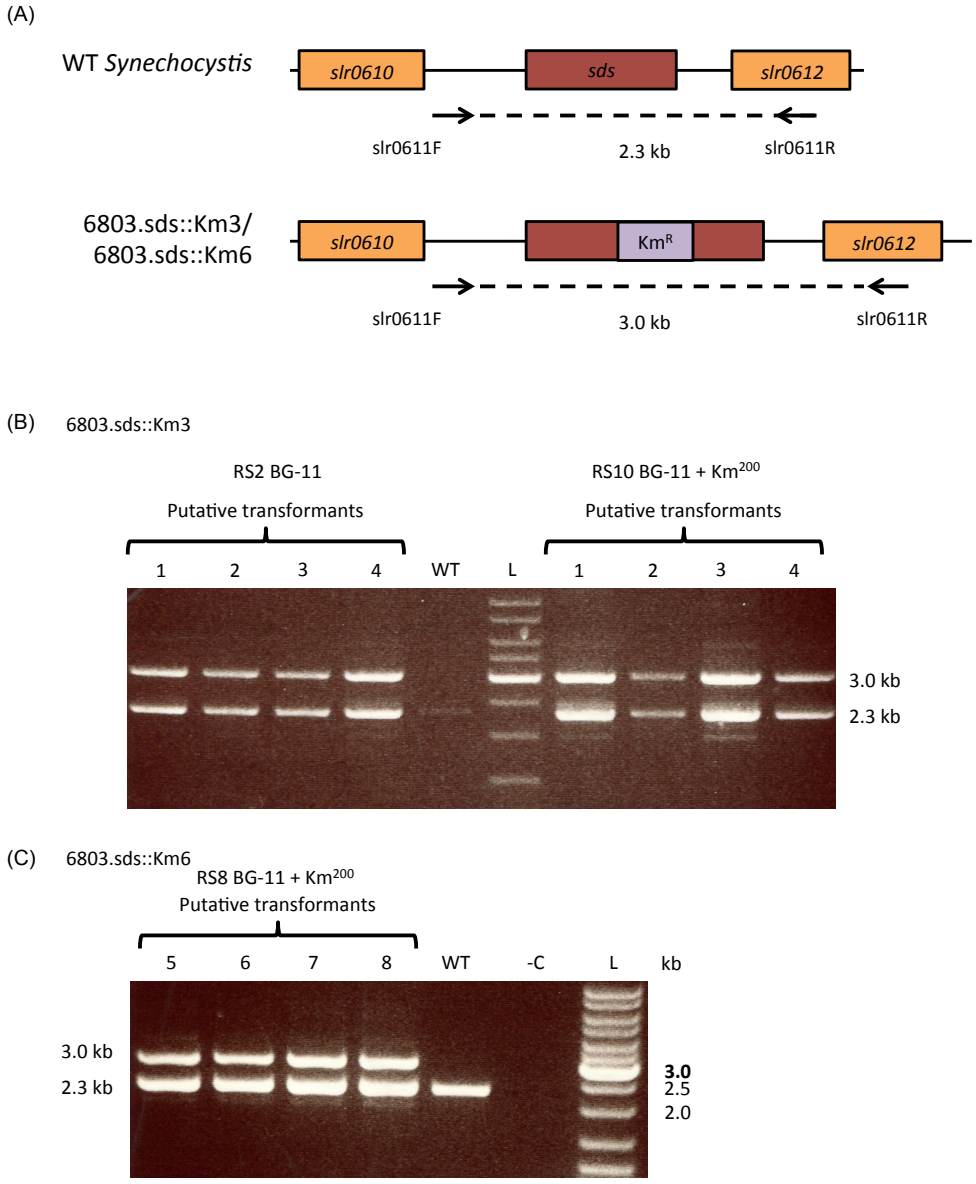


Figure 3.30 PCR screening of putative transformants of *Synechocystis* containing the *Km*^{R} cassette replacing the *sds* gene.

(A) Schematic showing the binding sites and expected fragments sizes for WT *Synechocystis* and successful 6803.sds::Km transformants. (B) PCR screening of putative transformants of 6803.sds::km3. WT *Synechocystis* (WT) was used a control. Heteroplasmic strains are present after restreaking 10 (RS10) times onto BG-11 media supplemented with kanamycin with two bands at 2.3 and 3 kb. (C) PCR screening of putative transformants of 6803.sds::km6. WT *Synechocystis* and no DNA (-C) was used a control. Heteroplasmic strains are still present after restreaking 8 (RS8) times onto BG-11 medium supplemented with kanamycin with two bands at 2.3 and 3 kb.

3.3 Discussion and future work

In summary, the research presented in this chapter demonstrates that exogenous addition of limonene is toxic to *Synechocystis* at concentrations as low as 0.02% (v/v) but using a dodecane overlay can mitigate this toxicity. The expression in *Synechocystis* of a synthetic gene encoding a plant LS was successfully achieved using three separate promoters (P_{nrsB}, P_{psbAII} and P_{cpc560}); however the production of limonene in these lines could not be demonstrated. Attempts were also made to delete the *sds* gene to increase the pool of GPP; however this gene was found to be essential.

During the course of this work, three separate reports were published (Davies et al., 2014; Halfmann et al., 2014b; Kiyota et al., 2014) that describe the successful production of limonene in three different engineered strains of cyanobacteria, including *Synechocystis*.

3.3.1 Toxicity tests

Both limonene enantiomers were toxic towards *Synechocystis* at concentrations as low as of 0.02% (v/v); however this toxicity can be mitigated by phase extraction with an organic layer of dodecane as limonene is poorly soluble in water but miscible in organic solvents. However since limonene was added exogenously, the results obtained may not be the same as when limonene is produced inside the cell, but the tests can give an indication of the effect limonene production may have on *Synechocystis*. The recovery of *Synechocystis* around 72 hours after limonene was added might be explained by the volatility of limonene. As the results from Figure 3.3 showed, after 24 hours some cells in the presence of limonene at 0.02, 0.1, and 0.2 % (v/v) survived and as limonene is known to be volatile, it may have evaporated and no longer be present at an inhibitory level in the culture.

In *E. coli* there is a correlation between toxicity of the solvent and its log$K_{O/W}$ value (Ramos et al., 2002). The octanol-water partition coefficient (reported as log$K_{O/W}$) of organic solvents often correlates with the antimicrobial activity of the solvent; therefore it is used as an indicator of solvent toxicity towards bacteria. The reported log$K_{O/W}$ for limonene was 4.40. If the value is between 1 and 5 then the product is expected to be hydrophobic and toxic towards the cell. This was the case for limonene towards *Synechocystis,* although the mechanism of toxicity is not clear. Hydrophobic aromatics like limonene are thought to accumulate within the cytoplasmic membrane, disrupting its integrity causing ions to leak, and a change in pH and in the proton gradient (Ramos et al., 2002). However, recently the toxicity of *d*-limonene towards *S. cerevisiae* was found to have an effect on the cell wall, not the cell membrane as previously thought (Brennan et al., 2013). The toxicity of limonene in *E. coli* was thought to be due to the common oxidation product limonene hyperperoxide

not limonene itself, and could be alleviated by a point mutation in an alkyl hyperoxidase (Chubukov et al., 2015).

Dodecane does not have a toxic effect towards *Synechocystis* and the toxicity of limonene is mitigated by the presence of a 20% (v/v) dodecane overlain on top of the culture. Dodecane has been used to trap the production of limonene in engineered strains of *Saccharomyces cerevisiae* and *E. coli* (Alonso-Gutierrez et al., 2013; Jongedijk et al., 2015). Dodecane also has been demonstrated to increase the yield and recovery of amorpha-4,11-diene, a volatile sesquiterpene, production in *E. coli* (Newman et al., 2006). The addition of the dodecane layer was used to increase the yield in limonene-producing strains of *Synechococcus* 7002, by trapping limonene that would most likely have been lost through evaporation (Davies et al., 2014). In *Synechococcus* 7002, the presence of a dodecane overlay did not interfere with the ability of atmospheric CO_2 to diffuse into the cultures (Davies et al., 2014). As the growth rate of 6803.cpc.LS was unaffected by the expression of LS, if LS is functional then the amount of limonene produced is not toxic to the cells. It might also be the case that the production of limonene inside the cells is not as toxic towards the cells as the exogenous addition of limonene.

3.3.2 Limonene production in cyanobacteria

Synechocystis strains expressing LS, under the control of various promoters, were successfully produced although synthesis of limonene could not be demonstrated. During the course of this work, the successful production of limonene in *Synechocystis* and two other cyanobacteria strains was published (Davies et al., 2014; Halfmann et al., 2014b; Kiyota et al., 2014). The differences in the strategies used are discussed below.

3.3.2.1 Extraction and GC-MS

The production of limonene was not detected in 6803.AII.LS strains after extraction by ethyl acetate and recovery using a dodecane layer for two-phase extraction. Both the pellet and supernatant after ethyl acetate extraction did not yield any detectable product and may not be the most suitable method for extracting limonene. As discussed previously the use of dodecane not only mitigates the possible toxic effect of limonene, but it can also be used as a method to trap limonene and has been used in strains of *E. coli*, *S. cerevisiae* and *Synechococcus* (Alonso-Gutierrez et al., 2013; Davies et al., 2014; Jongedijk et al., 2015). In *Synechococcus*, limonene was produced after the introduction of just the (-)-4*S*-LS gene from *Mentha spicata*, that was codon-optimised for *Synechococcus,* had the chloroplast transit peptide removed and was under the control of P_{cpc} promoter from *Synechocystis* (Davies et al., 2014). This method was not too dissimilar to the 6803.AII.LS

used for GC-MS analysis, although the *ls* gene was under the control of the weaker promoter, P_{psbAII}, and a different strain of cyanobacteria was used. To avoid the possibility of the reverse solvent effect, the dodecane from the overlaid sample was diluted by ethyl acetate. This was not done for the dodecane extract from the limonene producing *Synechococcus*, therefore if the dodecane overlaying 6803.AII.LS managed to trap any limonene, then it may have been below the level of detection by the GC-MS (Davies et al., 2014).

In *S. cerevisiae*, the detection of limonene was improved by trapping the volatile limonene produced in the headspace of the culture and more limonene was detected via this method than when dodecane was used (Jongedijk et al., 2015). Detection of limonene production was also achieved in *Anabaena* by trapping limonene in the culture headspace (Halfmann et al., 2014b). In limonene producing *Synechocystis*, limonene was detected in the cells using chloroform and from the cultures using a gas-stripping system to collect the volatile limonene (Kiyota et al., 2014). These alternative methods could also be used to try and trap volatile limonene produced from the *Synechocystis* strains produced in this chapter.

Other than the method of extraction, a lack of available substrate or a non-functional enzyme could also potentially explain why limonene was not detected in *Synechocystis*.

3.3.2.2 LS assay and detection using the DPPH assay

In the *ls* gene designed for insertion into *Synechocystis*, the codons for the threonine (T56) and glutamate (E57) residues just before RRX_8G motif were retained to provide stability to the enzyme. The *Mentha spicata ls* gene expressed was truncated for expression in *E. coli* just before RRX_8G and in *Synechococcus* 7002 before the glutamate residue before the RRX_8G motif (Alonso-Gutierrez et al., 2013; Davies et al., 2014). Truncated versions of *Mentha spicata* LS at the glutamate residue (E57) preceding the RRX_8W motif and at the glutamine (Q53) have been shown to be functional (Williams et al., 1998). The presence of the HA-tag at the C-terminus of LS could have an effect on the correct folding of the enzyme, although a $FLAG_2$-tag was fused to the C-terminus of the LS from *Picea sitchensis* that was successfully used to produce limonene in *Anabaena* sp. PCC 7120 (Halfmann et al., 2014b). This suggests that these design features are unlikely to render LS non-functional.

To see whether the enzyme expressed in 6803.cpc.LS was functional, an enzyme assay was performed using a DPPH assay; however, no limonene production was detected. As the assay was only attempted once, and no positive control was available, the lack of limonene detected may be due to flaws in the assay procedure rather than the non-function of LS. Significantly, the DPPH assay, was unable to detect limonene at less than 0.8 mM, which is

a high threshold. If this experiment were to be repeated, the dodecane extracted from the enzyme assay should be analysed by GC-MS as it is much more sensitive and would provide a clearer answer to whether the enzyme is functional.

3.3.3 Further metabolic engineering of *Synechocystis*

The unsuccessful attempt to replace all copies of the *sds* gene with a kanamycin resistance cassette, suggest that SDS is essential for *Synechocystis* and therefore its inactivation cannot be used to increase the pool of GPP. One of the main factors for choosing to delete *sds* was the claim on CyanoBase, that *sds* had been deleted and no mutant phenotype was seen. The results obtained in this chapter and further reading of the literature suggests that this is not the case. The solanesyl diphosphate is produced through the condensation reaction of isopentenyl diphosphate (IPP) to GPP or geranylgeranylpyrophosphate (GGPP) (Kanehisa and Goto, 2000). Solanesyl diphosphate forms the side chain of plastoquinone-9 (PQ-9) (Sadre et al., 2012). Further research into the literature indicated PQ-9 plays an essential role in the electron transport chain and although plastoquinones with shorter side chains can also be synthesised, *sds* appears to be essential to *Synechocystis* (Sadre et al., 2012).

An alternative strategy to increase the pool of GPP is to increase the flux to the pool rather than reduce the flux from the pool. In other microorganisms engineered to produce limonene, the pathway upstream of LS has been modified to increase the amount of GPP. For example, in several engineered strains a second copy of GPPS has been introduced (Alonso-Gutierrez et al., 2013; Halfmann et al., 2014b; Kiyota et al., 2014), and in *Synechocystis* an extra copy of the native genes found in the MEP pathway *dxs* (deoxyxylulose-1-phosphate synthase), *crtE* (GPP synthase) and *ipi* (isopentenyl pyrophosphate isomerase) were inserted to increase the yield of limonene (Kiyota et al., 2014). In *Anabaena*, the same enzymes were expressed to increase yield, however the native genes were not used; instead *dxs* from *E. coli*, *ipphp1* from *Haematococcus pluvialis* and GPPS from *Mycobacterium tuberculosis* were expressed (Halfmann et al., 2014b). If a lack of substrate was the reason why limonene was not produced to a detectable level, then the introduction of native/non-native genes from the MEP pathway or a non-native MVA pathway along with GPPS would all help to increase the flux to GPP and provide a larger substrate pool for LS. Finally, the introduction of an efflux pump has been shown to increase the yield of limonene in a microbial host, and is believed to function by relieving toxicity and preventing production inhibition (Dunlop et al., 2011).

Chapter 4 Molecular tools for trans-operon expression in *Synechocystis*

4.1 Introduction

For cyanobacteria to fulfil their potential as 'green' cell factories it is important to be able to control gene expression. In order to optimise the synthesis of biofuels and high-value chemicals in cyanobacteria the expression levels of each gene within a metabolic pathway need to be balanced to avoid the accumulation of any toxic products and remove any potential bottlenecks. The development of molecular tools that are able to tightly regulate the level of transcription and translation is vital to optimise a metabolic pathway for the production of novel compounds.

4.1.1 Transcriptional control tools - Promoters

In cyanobacteria, a range of native and foreign promoters have been used to express heterologous genes. Commonly, native promoters used to express trans-genes are taken from key photosynthetic genes e.g. P_{psbAII} (photosystem II D1 protein), P_{rbcL} (Rubisco large subunit), P_{cpc} (c-phycocyanin beta subunit) and P_{psaD} (photosystem I subunit II) (Wang et al., 2012). Inducible promoters are desirable when attempting to produce a toxic chemical, and one such endogenous promoter that has been exploited is the nickel-inducible promoter P_{nrsB} (Liu and Curtiss, 2009). Foreign promoters from *E. coli* (P_{lac}, P_{trc}/P_{tac}, P_{tet}) have also been used in various cyanobacteria, however they are not always effective in *Synechocystis* (Huang et al., 2010). Differences in the RNA polymerase holoenzyme between *E. coli* and *Synechocystis* may explain why *E. coli* promoters function differently in *Synechocystis* (Heidorn et al., 2011). Although, modifications made to certain *E. coli* promoters have helped improve their function in *Synechocystis* (Camsund et al., 2014; Huang et al., 2010).

4.1.1.1 P_{psbAII}

The P_{psbAII} promoter is considered to be a constitutive promoter, but higher levels of transcripts can be seen in high light conditions (Lindberg et al., 2010; Liu and Curtiss, 2009). P_{psbaII} is also a strong promoter and has been used to drive the production of ethanol, isoprene and fatty acids (Chen et al., 2014; Dexter and Fu, 2009; Lindberg et al., 2010).

4.1.1.2 Nickel-inducible promoter, P_{nrsB}

The nickel-inducible promoter, P_{nrsB}, originates from the Ni^{2+} sensing/response system in *Synechocystis* (García-Domínguez et al., 2000). The Ni^{2+} sensing/response system is made up of the *nrsBACD* operon and a two-component signal transduction system made up of *nrsR* and *nrsS*, found upstream of the operon. The *nrsBACD* operon is involved in Ni^{2+}

resistance whereas the *nrsR* and *nrsS* genes detect the presence of Ni^{2+} and induce the transcription of the operon. The presence of Ni^{2+} is thought to stimulate the kinase activity of NrsS, which phosphorylates NrsR. NrsR then binds to the intergenic region between *nrsRS* and *nrsBACD,* which activates the transcription of *nrsBACD* (López-Maury et al., 2002). P_{nrsB} has been used as a nickel inducible promoter to express heterologous genes. Liu and Curtiss expressed various lysin genes under P_{nrsB}, made up of the *nrsB* promoter, along with *nrsRS* to create a nickel-inducible lysis system in *Synechocystis* (Liu and Curtiss, 2009).

Within the Purton group the expression vector pLAH.nrsB (created by Lamya Al-Haj – see appendix) has been employed to express foreign genes under the nickel-inducible system. The P_{nrsB} in pLAH.nrsB, differs to the one used by Liu and Curtiss as it is made up of the *nrsB* promoter and *nrsR*, but does not contain *nrsS* (Liu and Curtiss, 2009). Nickel-induced expression of several genes under the P_{nrsB} has been demonstrated in *Synechocystis* (Al-Haj, 2014; Stoffels, 2014).

4.1.1.3 'Super-strong' promoter, P_{cpc560}

Recently, Zhou et al. published details of a 'super-strong' promoter, P_{cpc560} capable of producing functional proteins at very high expression levels (15% of total soluble proteins in *Synechocystis*) (Zhou et al., 2014). The P_{cpc560} is composed of two promoters and 14 transcription factor binding sites (TFBSs) found upstream of the *cpcB* gene encoding the beta subunit of c-phycocyanin and is responsible for expressing the *cpc* operon which encodes the phycocyanin and linker polypeptides present in phycobilisomes. The TFBSs region in the cpc560 promoter was demonstrated to be essential for high expression (Zhou et al., 2014). Such a strong promoter could be utilised in industrial biotechnology when using cyanobacterial for over-production of key enzymes or as a protein expression platform.

4.1.2 Translational control tools – ribosomal binding sites

As aforementioned in the introduction, translation in bacteria is initiated when the ribosome binds to mRNA at the RBS. The effectiveness of RBS is affected by the SD sequence, the presence of distance between the SD sequence and the start codon and the secondary structures that may develop due to the surrounding nucleotide sequence (Heidorn et al., 2011). In *Synechocystis*, an RBS sequence UAGUGGAGGU based on the optimal complementary sequence, based on the complementary base pairing to the 3' end of the 16S rRNA and with the ideal spacing between the central A and the start codon, was

compared against three BioBrick RBSs. Expression with the optimal *Synechocystis* RBS was greater than the BioBrick RBSs (Heidorn et al., 2011).

4.1.2.1 Intergenic regions containing RBS in operons

Operons are naturally found in bacteria and are a useful method of arranging related genes under the control of the same promoter. The presence of RBSs in the intergenic region (IG) between each gene can initiate translation of each gene in the operon. Although not much is known about RBSs in *Synechocystis*, genome analysis comparing the *Synechocystis* genome with other bacterial genomes indicated the IGs within an operon in *Synechocystis* were much longer (median distance between genes in a conserved pair was 56) than expected especially compared to other prokaryotes (Moreno-Hagelsieb and Collado-Vides, 2002; Price et al., 2005; Rogozin et al., 2002).

4.1.3 Background work

Within the group, two expression plasmids, pLAH.nrsB and pLAH.AII, are routinely used to express genes inserted into the genome at the *psbAII* locus, under the control of the P_{nrsB} and the P_{psbAII}, respectively. Another expression plasmid, pLAH.cpc, created by Aaron Lau, was produced to express genes at the *psbAII* locus under the P_{cpc560} promoter, designed by (Zhou et al., 2014). The pLAH.cpc expression vector was created to compare P_{cpc560} with P_{psbAII}. Two transgenic strains of *Synechocystis*, 6803.cpc.GFP and 6803.AII.GFP, were created by Aaron Lau with *gfpuv* expressed under the P_{cpc560} and P_{psbAII}, respectively (Figure 4.1).

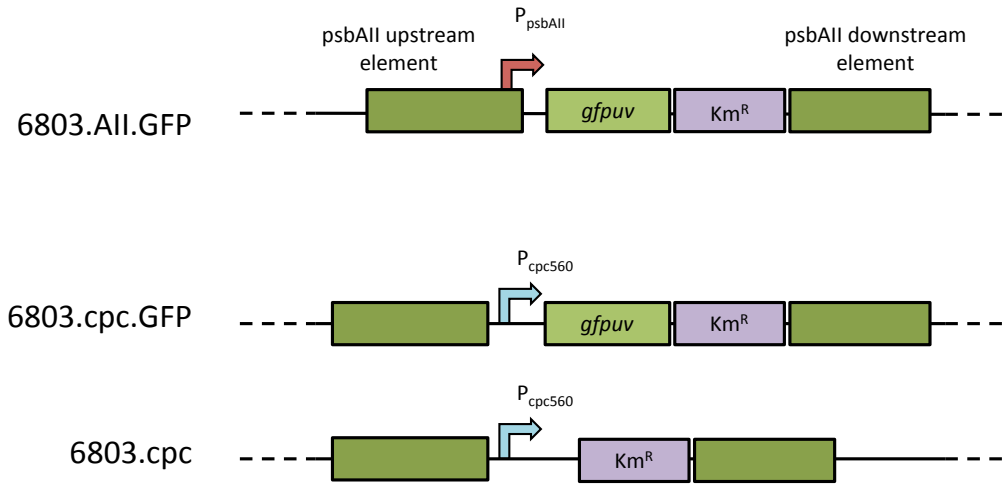


Figure 4.1 GFPuv expressing strains of *Synechocystis* and the corresponding control strain created by Aaron Lau.

The diagram shows the transgene arrangement in the strains created to test the expression of GFP under the control of the two promoters, P_{cpc560} and P_{psbAII}. A control strain, 6803.cpc was also created with the P_{cpc560} replacing the psbAII gene and promoter.

An expression plasmid termed "4op" was created by Ben Mackrow in the Purton group to express a four-gene operon (containing three antibiotic resistance genes and the gene for GFP) at the *Synechocystis psbAII* site under the control of P_{psbAII}. To design 4op, various native IGs thought to contain RBSs were used to separate the coding regions and ensure the translation of each gene in the operon. Short, native IGs were found within the *Synechocystis* genome by searching for sequences between known operons (using 'MicrobesOnline' database) (Alm et al., 2005). As the IGs would be incorporated into PCR primers used to assemble the operon, the size of the chosen IGs was limited to less than 40 bp. IG1, IG2 and IG3 were selected to separate the genes in 4op (Table 4.1). A transgenic strain (strain 6803.4op), containing the four-gene operon was created (Figure 4.2). In this strain expression of the first three genes was confirmed by phenotype testing, however expression of the fourth gene was not seen. A second expression plasmid, 4op.Rev, was created in which the order of the IG's was kept the same but the gene order was reverse (Figure 4.2). Again phenotype tests showed that the final gene was not expressed, suggesting that IG3 was not functional.

Table 4.1 IGs used in the 4op and 4op.Rev operons

The IGs used to separate the genes in the 4op and 4op.Rev operons created by Ben Mackrow. The location of the IG region is found between the two genes in the *Synechocystis* genome. The pOp value is the estimated probability that the gene pair are in the same operon, values that are closer to 1 or 0 are confident predictions that the two genes are in the same operon, whereas values near 0.5 are low-confidence predictions.

Name of Intergenic (IG) region	Location	Sequence	Length (bp)	pOp value
IG1	sll1030-sll1029	ttttccctataaattttgagcccaattaagcc	33	0.849
IG2	sll1032-sll1031	taattactagttgaccagccccccggattttgcc	35	0.919
IG3	sll1324-sll1323	gtgccctatagatctagttgccatt	26	0.993

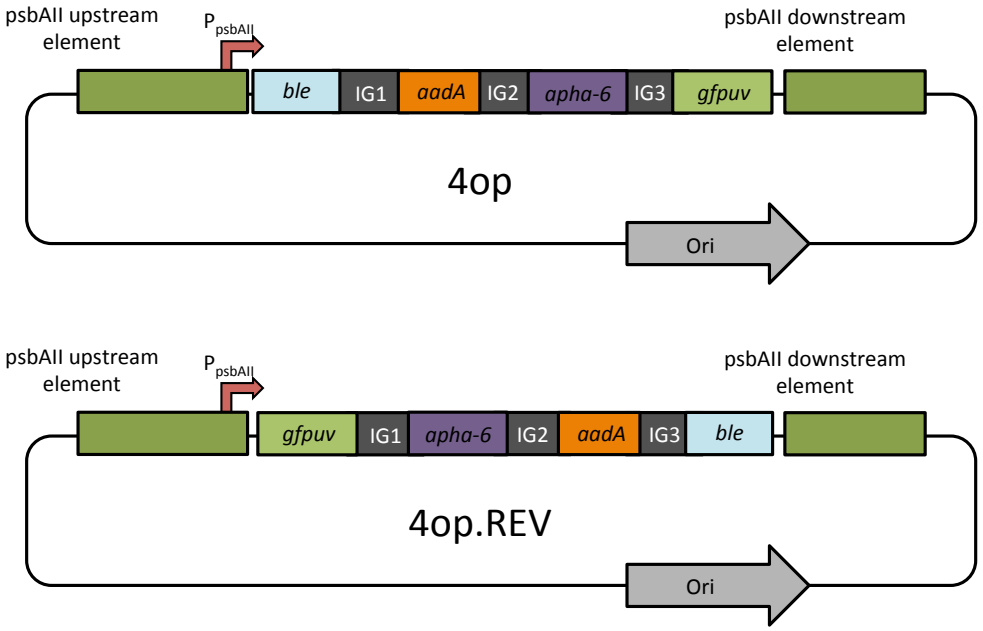


Figure 4.2 Expression plasmids used to create the transgenic strains of *Synechocystis* expressing a 4-gene operon under the control of P_{psbAII}.

The diagram shows 4op, the expression plasmid used to create the strain 6803.4op. The 4-gene operon shows the antibiotic resistance genes *ble*, *aadA* and *aphA-6*, which confer resistance to zeocin, spectinomycin and kanamycin, and *gfpuv* separated by IG1, IG2 and IG3, under the control of P_{psbAII}. Also shown is the expression plasmid 4op.Rev (gene order reversed relative to 4op), used to create the strain 6803.4op.Rev.

4.1.4 Aims and objectives

The overall aim of the work presented in this chapter was to develop further the necessary tools for efficient trans-operon expression in *Synechocystis*.

This builds on the work previously carried out in the lab. The expression of the multi-gene operon (designed by Ben Mackrow) was tested under the control of the nickel inducible promoter, to investigate whether an inducible promoter such as P_{nrsB} can regulate individual genes within an operon. A further promoter, P_{cpc560} was studied and compared with P_{psbAII} using the transgenic strains 6803.cpc.GFP and 6803.AII.GFP (created by Aaron Lau). Finally, continuing on from the work started by Ben Mackrow, more intergenic regions containing potential RBSs were found and tested within a two-gene operon that used as GFP as the reporter, to test the functionality of each intergenic region.

4.2 Results

4.2.1 Expression of the multi-gene operon under *nrsB* promoter

To determine whether an inducible trans-operon system could be established for *Synechocystis*, part of the multi-gene operon in 4op was used to create a transgenic strain containing a multi-gene operon under the control of the native nickel inducible promoter, P_{nrsB}. The operon was composed of three antibiotic resistance genes: *ble*, *aadA* and *aphA-6*, that confer resistance to zeocin, spectinomycin and kanamycin (zsk), respectively and is therefore referred to as the zsk operon. Intergenic regions IG1 and IG2 were used to separate the genes in the operon.

4.2.1.1 Creation of transgenic strain expressing the zsk operon under the nickel-inducible promoter

To create the expression plasmid, pLAH.nrsB.zsk (Figure 4.3) used to transform *Synechocystis*; the *ble*-IG1-*aadA*-IG2-*aphA-6* region was amplified from 4op and cloned into a modified pLAH.nrsB expression vector, pLAH.nrsB_cat. The pLAH.nrsB_cat plasmid was created in order to remove the Km^R cassette present between the *psbAII* elements and insert a *cat*, a chloramphenicol resistance gene, into the backbone to enable selection in *E. coli*.

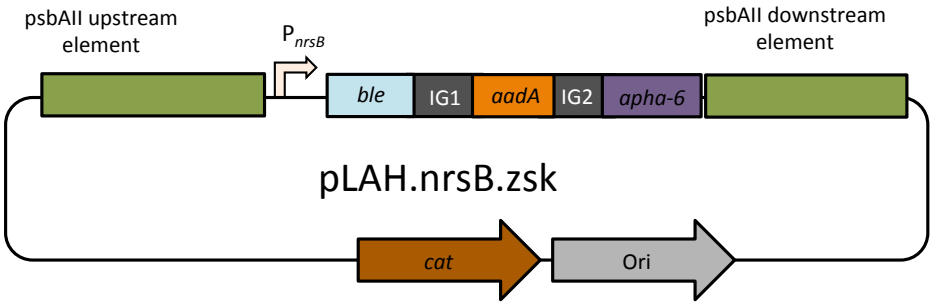


Figure 4.3 Diagram of the expression plasmid used to create transgenic strains of *Synechocystis* expressing the zsk operon under the P_{nrsB} at the *psbAII* site in the genome.

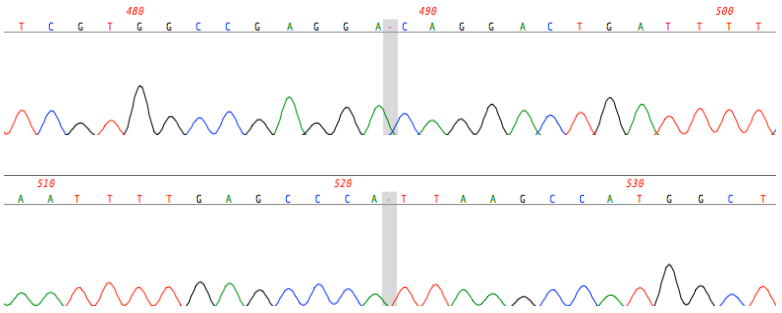
The zsk operon was ligated into the pLAH.nrsB_cat vector in order to create pLAH.nrsB.zsk.

The region was amplified using AL.ZnR.4op.F and AL.KanR.4op.R, and produced NdeI and Acc65I restriction sites at the 5' and 3' end of the PCR product, respectively. The PCR fragment containing the multi-gene operon was initially cloned into the pJET1.2/blunt vector to create pJET.zsk, and then excised using NdeI and Acc65I and cloned in the NdeI and Acc65I site of pLAHnrsB_cat to create pLAH.nrsB.zsk. Restriction analysis and sequencing using primers nrsB.seqF and Out.GFP.R verified the construction of

pLAH.nrsB.zsk. However, when sequenced with nrsB.seqF, a single base deletion is observed towards the end of the *ble* gene, leading to a frame shift (Table 4.2). Since the growth of the cell line was seen in the presence of Zeocin, the mutation does not appear to have an effect on the function of the Ble protein. These results also showed a single base deletion in IG1, in the region where the SD sequence in likely to be present this may have an effect on the expression levels of *aadA* (Table 4.2).

Table 4.2 Deletion mutations in *ble* and IG1 present in pLAH.nrsB.zsk

Table shows the expected and actual sequence of the end of *ble* (in blue) and IG1 (in black) and the start of GFP (in green). The expected and actual amino acid sequence of the C-terminal end of Ble is also shown. The chromatogram received from sequencing pLAH.nrsB.zsk with nrsB.seqF shows the missing bases in *ble* and IG1

Expected pLAH.nrsB.zsk sequence	…GTGGCCGAGGAGCAGGACTGAttttccctataaattttgagcccaattaagccATGGCTCGTGAAG…
Expected amino acid sequence	… V A E E Q D *
Consensus Sesquence	…GTGGCCGAGGA-CAGGACTGAttttccctataaattttgagccca-ttaagccATGGCTCGTGAAG…
Actual pLAH.nrsB.zsk sequence	…GTGGCCGAGGACAGGACTGAttttccctataaattttgagcccattaagccATGGCTCGTGAAG…
Actual amino acid sequence	… V A E D R T D F P Y K F *
Chromatogram of nrsB.seq.F on pJAH.nrsB.zsk	

WT *Synechocystis* was transformed with pLAH.nrsB.zsk, to create a transgenic strain with the zsk operon. To select for transformants, the cells were supplemented with 6.4 μM Ni^{2+} to induce the expression of the antibiotic resistance genes. Colonies were selected on $Zeo^{7.5}$ + $Spec^{30}$ + Km^{30} for resistance to all three antibiotics. 16 colonies were restreaked onto plates containing $Zeo^{7.5}$ + $Spec^{30}$ + Km^{30}+ 6.4 μM Ni^{2+}. Of the 16 colonies, 10 grew and PCR using Ben.ups.F and C.Frag.BR primers was performed on the genomic DNA extracted from these putative transformants. A single transformant (6803.nrsB.zsk_7) appeared to have the operon integrated at the *psbAII* locus (Figure 4.4). A further 16 putative transformants were tested, and from this two additional transformants (6803.nrsB.zsk_D and 6803.nrsB.zsk_M) appeared to have successfully integrated the operon at the *psbAII*

site (Figure 4.4). All successful transformants were judged to be homoplasmic based on the absence of the smaller PCR product from the WT locus (Figure 4.4). The unusually high number of false transformants seen in this transformation may be due to the low antibiotic concentrations used, which may have enabled so many strains to become resistant and grow.

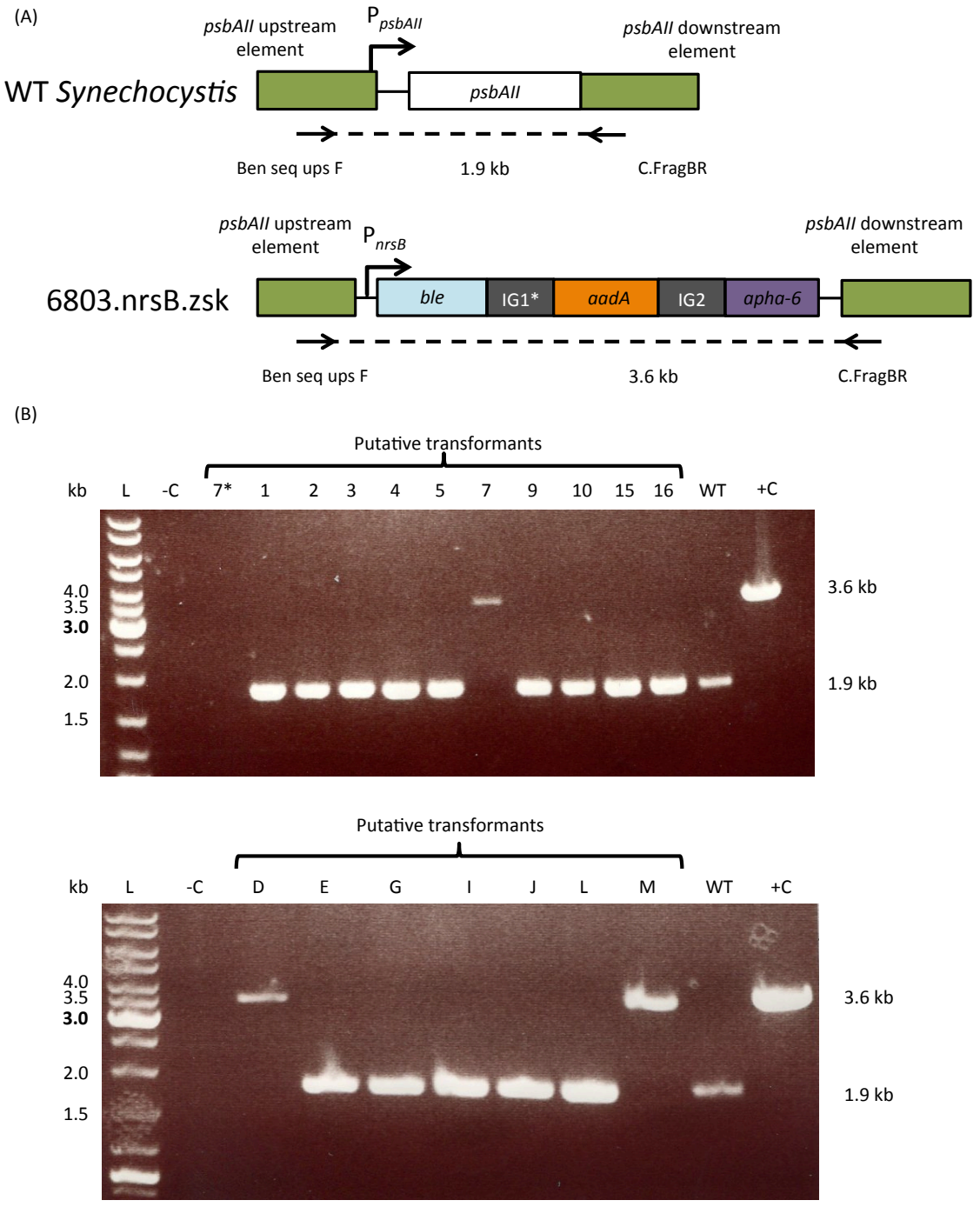


Figure 4.4 PCR screen of successful transformants in *Synechocystis* containing the zsk operon under the control of P_{nrsB}

(A) Schematic showing the binding sites and expected fragment sizes for WT *Synechocystis* and successful 6803.nrsB.zsk transformants. (B) PCR screens of putative transformants of 6803.nrsB.zsk. WT *Synechocystis* (WT), pLAH.nrsB.zsk (+C) and no DNA (-C) were used as controls. Successful transformants 7, D and M all have a band at 3.6 kb and are homoplasmic.

4.2.1.2 Testing the expression of the antibiotic resistance genes within the operon

A growth 'spot' test was performed to assess whether the nickel-inducible promoter was able to control the expression of the zsk operon such that the transformants would show sensitivity to the antibiotics in the absence of added nickel, but resistance when nickel was added to the medium.

The three biological replicates of 6803.nrsB.zsk were grown, alongside WT *Synechocystis* and 6803.4op (used as controls) on BG-11 medium containing zeocin/ spectinomycin/ kanamycin and with/without 6.4 μM Ni^{2+} (Figure 4.5). After four days, the 6803.nrsB.zsk transformants 7 and D showed a clear phenotype of growth on zeocin and spectinomycin in the presence of nickel, but not in its absence. This demonstrates that the expression of *ble* and *aadA* are nickel–induced and under the control of P_{nrsB}. For transformant M, the induction appears much weaker suggesting a possible mutation in the promoter element. However, all three 6803.nrsB.zsk transformants were able to grow in the presence of kanamycin even when Ni^{2+} was absent (Figure 4.5). The expression of *aphA-6* is therefore not induced by Ni^{2+}, and is not under the control of P_{nrsB}. A possible explanation for this constitutive expression is given in section 4.2.1.3.

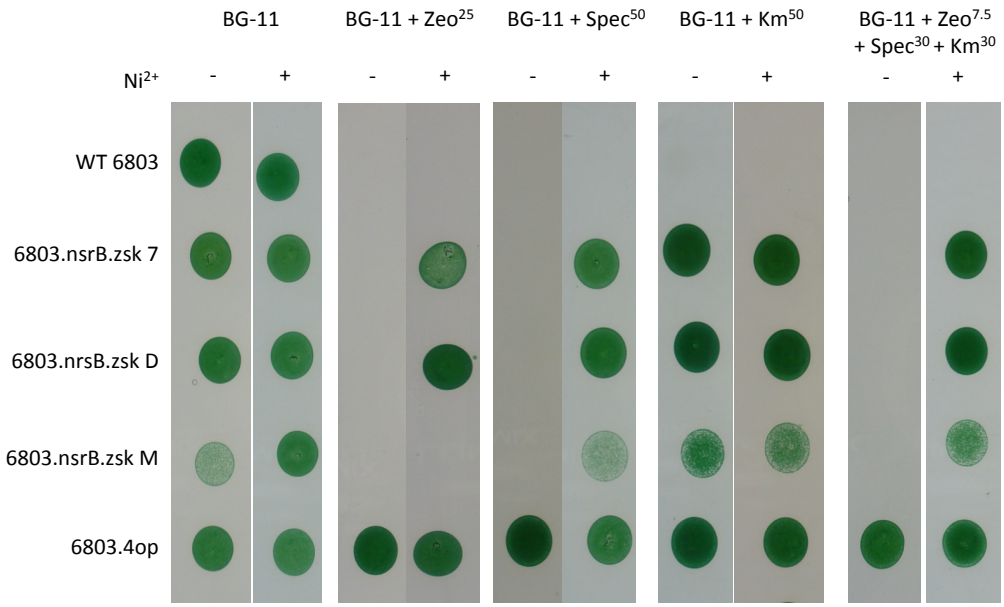


Figure 4.5 Growth 'spot' test assessing the expression of the zsk operon in 6803.nrsB.zsk transformants in the absence (−) and presence (+) of Ni^{2+}.

Photo taken after 4 days. Three 6803.nrsB.zsk transformants were grown alongside WT *Synechocystis* (WT 6803) and 6803.4op (used as the controls). WT 6803 was unable to grow on any of the antibiotics and grew only on BG-11 as expected whereas 6803.4op (expressing *ble*, *aadA* and *aphA-6* under $P_{psbAII)}$ was able to grow on all plates. Growth of 6803.nrsB.zsk on zeocin and spectinomycin was induced by the presence of Ni^{2+} but 6803.nrsB.zsk was able to grow on kanamycin even when Ni^{2+} was absent.

As 6803.nrsB.zsk is resistant to spectinomycin, this suggests that the intergenic region IG1* (IG1 with the 1bp deletion), is functional allowing translation of the *aadA* coding region. However, after 7 days, all three 6803.nrsB.zsk started to grow on spectinomycin, even when Ni^{2+} was absent (Figure 4.6). This probably reflects leaky expression from P_{nrsB} either due to the low concentration of zinc in the medium or the very weak constitutive promoter in P_{nrsB} (further examined in the discussion).

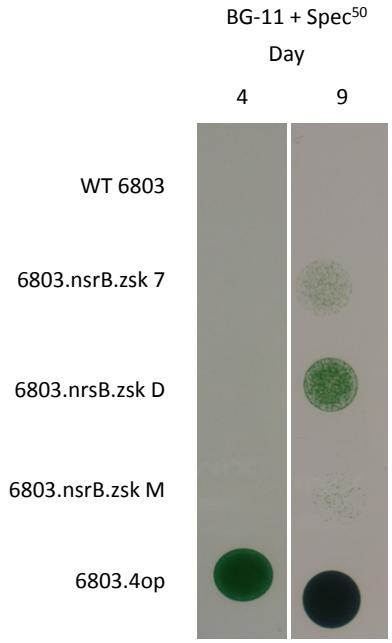


Figure 4.6 Growth 'spot' test assessing the expression of *aadA* in 6803.nrsB.zsk transformants.

Photos taken 4 and 9 days after 'spotting' onto medium containing spectinomycin. Three 6803.nrsB.zsk transformants were grown alongside WT *Synechocystis* (WT 6803) and 6803.4op (used as the controls). WT 6803 was unable to grow on plates containing spectinomycin as expected and 6803.4op (expressing *aadA* under P_{psbAII}) was able to grow in the presence of spectinomycin. The 6803.nrsB.zsk transformants showed weak growth after 9 days.

4.2.1.3 Investigating the presence of a possible cryptic promoter

As the expression of *aphA-6* occurs readily in the absence of nickel, a cryptic promoter could be present in *aadA* or IG2 that enables expression.

Using the *E. coli* promoter prediction software, BPROM the *aadA* gene and IG2 upstream of *aphA-6* was analysed and two predicted promoters were found. One promoter was predicted to be in the middle of the spectinomycin resistance gene and the other one towards the end of the *aadA*, with the predicted -10 box region formed from both *aadA* and IG2 (Figure 4.7). The finding that such cryptic promoters can confound the intended phenotype highlights one of the challenges of synthetic biology – namely, the difficulty of predicting and avoiding such 'off-target' effects from unintended *cis* elements within transgene constructs. Nonetheless, the expression of AphA-6 demonstrates that IG2 is also functional in terms of mediating efficient translation.

ATGGCTCGTGAAGCGGTTATCGCCGAAGTATCAACTCAACTATCAGAGGTAGTTGGCGTCAT
CGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCG
GCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATGAAACA
ACGCGGCGAGCTTTGATCAACGACCTTTTGGAAACTTCGGCTTCCCCTGGAGAGAGCGAGAT
TCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAG
CTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAG
CCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTGC
CTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGG
CGCTAAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGCGATGAGCGAAAT
GTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGA
TGTCGCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAG
CTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAA
GAATTTGTCCACTACGTGAAAGGCGAGATCACTAAGGTAGTTGGCAAATAATAATTACTAGT
TGACCAGCCCCCCGGATTTTGCCATGACCATGGAATTACCAAAT

aadA
apha-6
IG2
–10 box
–35 box

Figure 4.7 Sequence of the end of *aadA*, IG2, and the starts of *aphA-6* in 6803.nrsB.zsk.

The –10 box and –35 box of the cryptic promoter predicted to be in the middle of the *aadA* gene and at the end of the gene are underlined.

4.2.2 Expression of GFP under the *cpc* promoter

To compare the expression of genes under the P_{cpc560} and P_{psbAII} promoter elements, and test the claims of P_{cpc560} being a super-strong promoter, the expression of *gfp* was tested in the two transgenic strains 6803.cpc.GFP and 6803.AII.GFP. Western blot analysis and the fluorescence of GFPuv were used to compare the expression of GFPuv under the two promoters.

Comparing the expression of *GFPuv* under the control of the *psbAII* promoter and the *cpc560* promoter.

To compare the expression of GFPuv under the two promoters the following cultures were grown in BG-11: three biological replicates of 6803.cpc.GFP (#1, 5, 7) and 6803.AII.GFP (#1, 2, 7) with 6803.cpc and 6803.AII (see appendix) as the two control strains. Once the cultures were grown to an OD_{730} of ~0.9, the cells were used for Western blot analysis and fluorescence measurements.

For Western blot analysis, samples were probed with anti-GFP antibodies and anti-RbcL antibodies as a loading control, and the Odyssey Infrared Imaging System was used for the quantification of GFP protein levels. The expression of *gfp* was seen in all 6803.cpc.GFP

and 6803.AII.GFP lines with a band at 27 kDa. This band is absent in both control strains (Figure 4.8A). Western blot confirms expression of *gfp* under both P_{cpc560} and P_{psbAII} promoters.

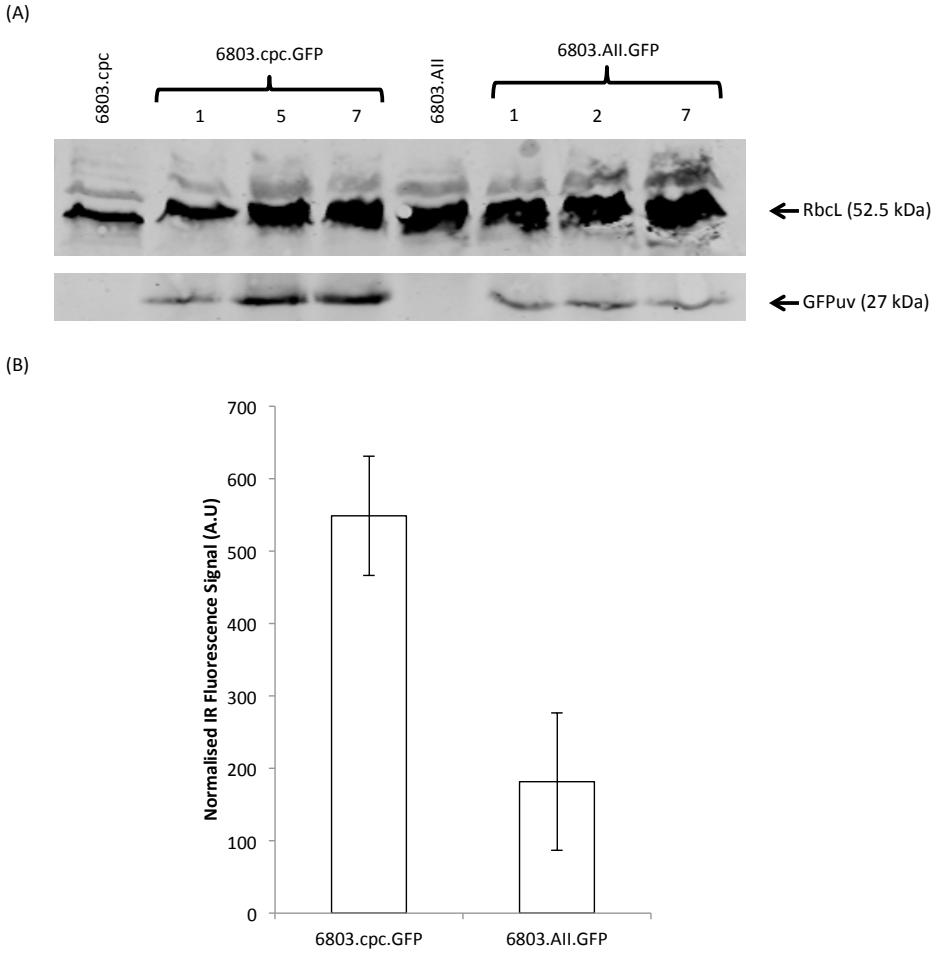


Figure 4.8 Western blot analysis confirms and compares GFPuv expression under P_{cpc560} and P_{psbAII} in transformants of *Synechocystis*.

(A) Western blot analysis was performed with whole cell extracts and anti-GFP and anti-RbcL antibodies, using the Odyssey® Infrared Imaging System for detection. Samples were run on a 12% SDS-PAGE gel. The expected band size for GFP is 27 kDa and 52.5 kDa for RbcL, which acted as loading control. Protein size was determined using the PageRuler™ prestained protein ladder (Thermo Scientific). Strains 6803.AII and 6803.cpc acted as negative controls. (B) The GFPuv signal was quantified using the Odyssey® Infrared Imaging System. The GFPuv IR fluorescence signal was normalised against the signal of RbcL. The average normalised IR signal for all the three 6803.cpc.GFP strains was compared to 6803.AII.GFP. Error bars are ± one SD.

The IR fluorescence signal of GFPuv and RbcL was quantified using the Odyssey Infrared Imaging System. As seen in Figure 4.8 the level of GFPuv appears to be at least two-fold greater under the control of P_{cpc560} than P_{psbAII}, supporting the reported finding that P_{cpc560} can serve as a very strong promoter (Zhou et al., 2014).

GFP fluorescence was measured using a luminescence spectrometer. However, the emission maximum for GFP at 509 nm is obscured by endogenous emissions with a peak at 455 nm that arises from the light-harvesting pigments present in cyanobacteria (Figure 4.9) (Spence et al., 2003).

In order to remove the effect of the peak at 455nm, the absorbance units for each wavelength was normalised to the absorbance value at 455 nm. The spectrum of the control strain was then subtracted from the test strain and plotted. The peaks of 6803.cpc.GFP_5 and 6803.cpc.GFP_7 are higher than the 6803.AII.GFP transformants. 6803.cpc.GFP_1 appears to be an anomaly, with a small peak in comparison to the others transformants. Ignoring 6803.cpc.GFP_1, the results using the luminescence spectrometer match the results of the Western blot analysis; the expression of GFP is greater under the control of P_{cpc560} (Figure 4.10).

Figure removed

Figure 4.9 The absorbance spectra of *Synechocystis* (black line).

The green line represents the emission peak at approximately 509nm of GFP, which overlaps with part of the absorption peak seen in *Synechocystis*. This absorption of *Synechocystis* may interfere with detection of GFP. Modified image from http://2010.igem.org/Project_usu.html

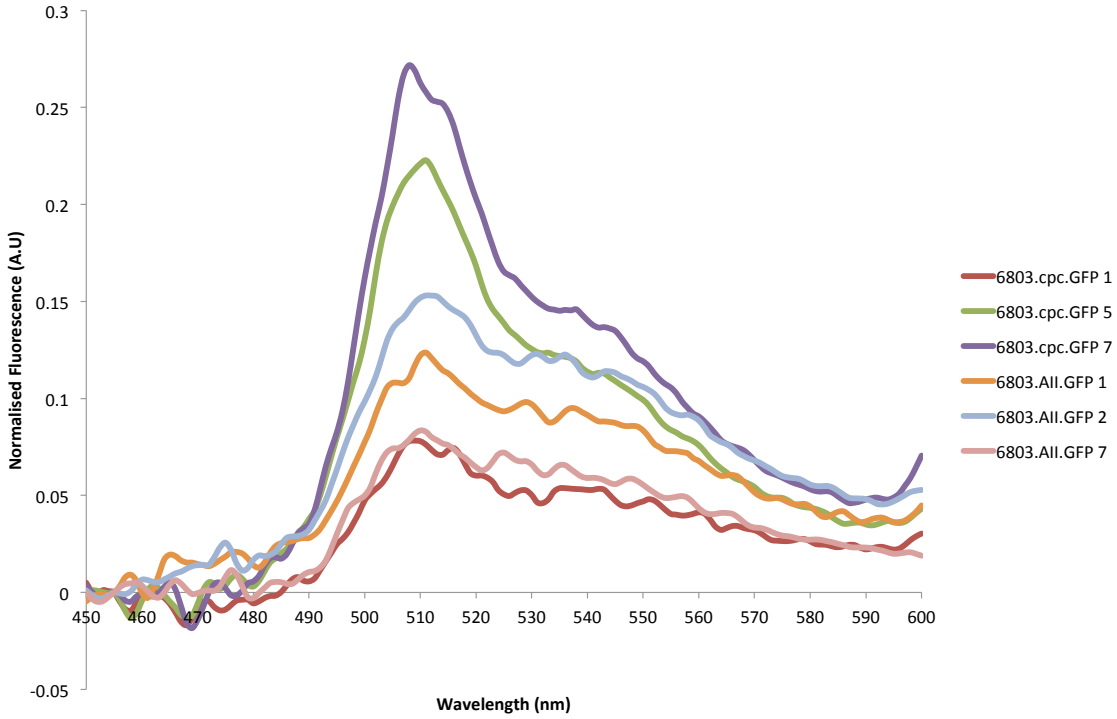


Figure 4.10 GFPuv Fluorescence measured under the control of promoter, P_{cpc560} and P_{psbAII} in transformants of *Synechocystis*.

Cultures of similar OD_{730} were excited at 395 nm and the emission spectrum between 450 – 600 nm was plotted. The absorbance of each wavelength was normalised to the absorbance value at 455 nm and the spectrum of the control strain (6803.cpc or 6803.AII) was then subtracted from these values. The emission peak at 509 nm is GFPuv.

4.2.3 Testing intergenic regions (containing RBS) in a two-gene operon

Continuing from the work carried out by Ben Mackrow in the Purton group, more IGs (containing RBS) for insertion in an operon were identified from the genome of WT *Synechocystis*. Various IGs were then tested by placing these between the coding regions of *ble* and *gfpuv,* in a two-gene operon, under the control of P_{cpc560} at the *psbAII* site (Figure 4.11). The expression of *gfpuv* was used to check the functionality of the IGs.

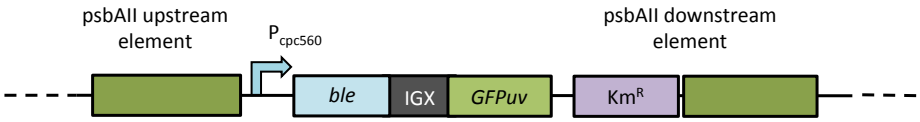


Figure 4.11 Two-gene operon created to test various IGs.

Diagram shows the two-gene operon, made up of *ble* and *gfpuv* separated by the IG, to test the function of the IG.

4.2.3.1 Search and identification of IGs to test

As IG2 (Figure 4.5) has already been shown to be functional, IG2 was chosen for use in the two-gene operon. In addition, a non-native IG region, IG4, was designed that utilises the optimal RBS, RBS* designed by Heidorn et al. (2011), and extra bases were added to make the IG 22 bases in length. More native IGs were initially identified by searching the 'Microbes Online' database' for pairs of genes that were predicted to lie within the same operon and had a pOp value (estimated probability that the pair were in the same operon) that was close to 1 or 0 as this gives more confident predictions that the two genes are within an operon, whereas values near 0.5 are low-confidence predictions. The IG sequences between the coding pairs were then chosen for pairs that were also less than 40 bases apart using CyanoBase (Table 4.3). From the list of IGs, the IG between *sll0939* and *sll0938* was renamed IG5 and was chosen for testing.

Table 4.3 List of the native intergenic regions found in *Synechocystis* genome that were between 20 and 40 bases.

The sequence for the IG was obtained for each gene pair. The function of the gene products in the gene pairs was found using CyanoBase. Highlighted in blue is IG5, which was chosen for testing.

Gene Pair	Function of Genes	Length (bp)	Sequence 5' to 3'	pOp value
sll1204/ sll1203	similar to macrolide efflux protein/ hypothetical protein	17	TTAATTTACAAAAATAA	0.965
sll1062/ sll1061	unknown protein/ unknown protein	26	ACCCTCTTAAAAAAATCTAAAAAACT	0.956
sll0710/ sll0709	unknown protein/ putative endonuclease	40	TCTAAGTTTTCTTTTCATTGAATAGTTTTATTTACATATT	0.922
slr0722/ slr0723	hypothetical protein/ hypothetical protein	20	TTTTTTTGGCCCTTTTCCCT	0.938
slr0250/ slr0251	hypothetical protein/ ATP-binding protein of ABC transporter	24	GCTGTGGGATACTAACGGCATTCG	0.952
sll0218/ sll0217	hypothetical protein/ flavoprotein	26	GTTCACCTTCTACCTGGGATTTTCTC	0.962
slr1103/ slr1104	hypothetical protein/ hypothetical protein	28	CTTAGAATTAATAGATCCCAGATTCTCT	0.911
sll1680/ sll1679	hypothetical protein/periplasmic protease HhoA	33	GTTTCCCGTACCTCACCCATCCAAACTTTAACC	0.844
sll0939/ sll0938	hypothetical protein/aspartate transaminase	24	ACTAAAAGGCAAAGGAATAAAACT	0.772
sll0933/ ssl1784	hypothetical protein/ 30S ribosomal protein S15	17	GTTTTTGCATTTTTCCC	0.967
slr0984/ slr0985	CDP-glucose 4,6-dehydratase/dTDP-4-dehydrorhamnose 3,5-epimerase	18	AACAAAAACCCAGCAAAA	0.966
slr1617/ slr1618	similar to UDP-glucose 4-epimerase/ unknown protein	12	GAAAGCAAGCAA	0.847
slr1618/ slr1619	unknown protein/ hypothetical protein	28	TTAACTTATTCAATTAAGAGGTAATTCA	0.75
slr1063/ slr1064	probable glycosyltransferase/ probable glycosyltransferase	21	AACTAGTTTACGTTTCATAAT	0.928

In order to test whether an IG was actually required for translation of a downstream cistron, the expression of *gfpuv* was also tested without the addition of any IG so the start codon of *gfpuv* was placed immediately after the *ble* stop codon (construct named IG0bp). Similarly, the expression was also tested with a single base between the stop codon of *ble* and the start codon of *gfpuv* (named IG1bp).

A list of the IGs used to test in the two-gene operon is given in Table 4.4.

Table 4.4 List of the IGs tested in the two-gene operon containing *ble* and *gfpuv*

IG regions to be tested, both IG2 and IG5 are native IG regions. IG4 has been designed to complement the 16S rRNA. IG0bp represent no IG between the stop codon of the first gene and the start codon of the second gene. With IG1bp there is a single base between the stop codon of the first gene and the start codon of the second gene. The pOp value is the estimated probability that the gene pairs are in the same operon.

Name of Intergenic (IG) region	Gene pair	Sequence 5'-3'	Length (bp)	pOp value
IG2	sll1032-sll1031	TAATTACTAGTTGACCAGCCCCCCGGATTTTGCC	35	0.919
IG4	N/A	TGGCAATAGTGGAGGTAATCGA	22	N/A
IG5	sll0939-sll0938	ACTAAAAGGCAAAGGAATAAAACT	24	0.772
IG0bp	N/A	N/A	0	N/A
IG1bp	N/A	G	1	N/A

4.2.3.2 Creation of transgenic strains expressing the two-gene operon

To construct the expression plasmids required to create the transgenic strains, first *ble* and *gfp* were individually amplified to each have the IG region in place. The PCR products (ble_IGX and IGX_GFP) were then used as a template to perform an overlap PCR to join the two genes together (ble_IGX_GFP). The resulting PCR product was then cloned into pLAH.cpc, to produce pLAH.ble_IGX_GFP, (for further details see 2.8.1). Figure 4.12 demonstrates the creation of the transgenic strain containing the two-gene operon, when IG2 is used, but the same method was used for all the strains tested.

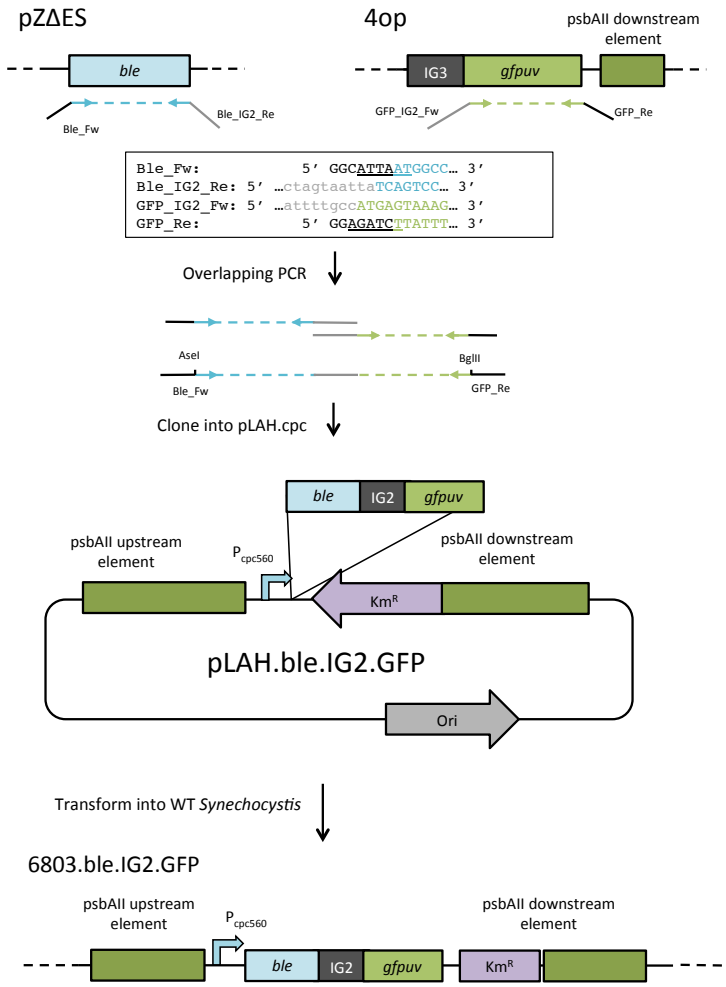


Figure 4.12 Schematic showing the construction of transgenic strains expressing the two-gene operon

Example showing the construction of 6803.ble.IG2.GFP, but the process applies for all strains. The primers used to amplify *ble* from pZΔES (See appendix) were designed so that the forward primer would have an AseI restriction site at the 5' end and the reverse primer designed to have the entire intergenic region after the stop codon of *ble*. To amplify *gfpuv* from 4op, the forward primer was designed so the entire intergenic region is upstream of the start codon of *gfpuv* and the reverse primer was designed to have a BglII restriction site after the stop codon. This was the case for when the intergenic region was IG2, IG4 and IG5, but when the intergenic region was 0bp or 1bp, the reverse primer for amplifying *ble* was designed so that after the end of *ble* the start of *gfpuv* was present after the 0bp or 1 bp (G) space after *ble*. The forward primer for amplifying *gfpuv* was designed so that before the start of *gfpuv*, the end of *ble* was present before the 0bp or 1 bp (G) space.

WT *Synechocystis* was then transformed with the various pLAH.ble.IGX.GFP expression plasmids to create the transgenic lines. After selection on kanamycin supplemented media, six putative transformants of 6803.ble.IG2.GFP, 6803.ble.IG4.GFP and 6803.ble.IG5.GFP were tested. PCR performed with Ben.ups.F and C.Frag.BR, on genomic DNA, confirmed the successful integration of the two-gene operon at the *psbAII* site (Figure 4.13). PCR was also performed on three putative transformants of 6803.ble_IG0bp_GFP and

6803.ble_IG1bp_GFP, and confirmed the successful integration of the two-gene operon at the *psbAII* site to create 6803.ble.IG0bp.GFP and 6803.ble.IG1bp.GFP (Figure 4.14). All 6803.ble.IG2.GFP, 6803.ble.IG4.GFP, 6803.ble.IG5.GFP, 6803.ble.IG0bp.GFP and 6803.ble.IG1bp.GFP strains that showed a transformant band were homoplasmic.

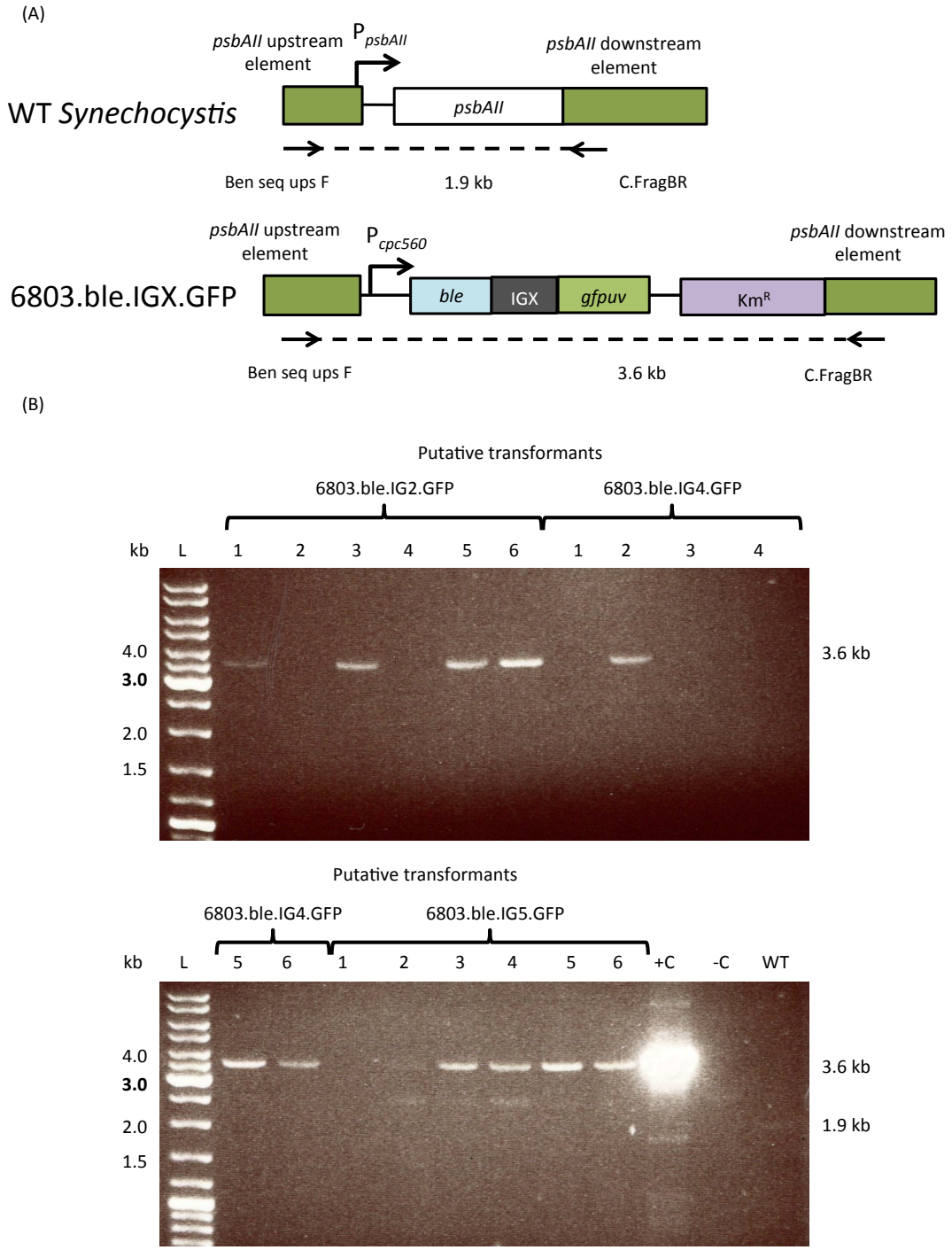


Figure 4.13 PCR screen of successful transformants in *Synechocystis* containing the two-gene operon with the various IGs (IG2, IG4, IG5) under the control of P_{cpc560}

(A) Schematic showing the binding sites and expected fragment sizes for WT *Synechocystis* and successful 6803.ble.IGX.GFP strains. (B) PCR screens of putative transformants of 6803.ble.IGX.GFP strains. WT Synechocystis (WT), no DNA (-C) and pLAH.ble.IG2.GFP (+C), expected to give a band at 3.6 kb, were used as controls. Transformants with a band at 3.6 kb and no visible band at 1.9 kb are presumed to be homoplasmic.

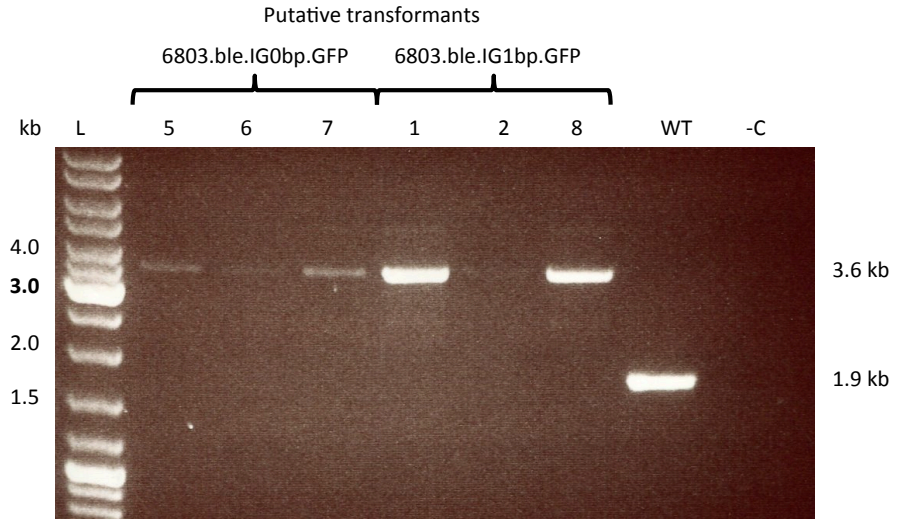


Figure 4.14 PCR screen of successful transformants in *Synechocystis* containing the two-gene operon with either IG0bp or IG1bp.

Colony PCR screens of putative transformants of 6803.ble.IGX.GFP strains. WT *Synechocystis*, WT and no DNA template (-C), were used as controls.

4.2.3.3 Testing GFP expression and IG function

As in the preceding studies, Western blot analysis and fluorescence of GFPuv were used to compare the expression of *gfpuv* when linked to the various IGs.

Three biological replicates of each 6803.ble_IG2_GFP, 6803.ble_IG4_GFP and 6803.ble_IG5_GFP were grown alongside 6803.cpc and 6803.cpc.GFP, used as negative and positive controls respectively, to an OD_{730} of ~0.7. For Western blot analysis samples were probed with anti-GFP and anti-RbcL antibodies (loading control) and the Odyssey Infrared Imaging System was used for the quantification of GFP (Figure 4.15). The preliminary results confirmed the expression of GFP with a band at 27 kDa in all three 6803.ble_IG2_GFP, 6803.ble_IG4_GFP and 6803.ble_IG5_GFP strains. The band is present in the positive control 6803.cpc and absent in 6803.cpc as expected. IG2, IG4 and IG5 are therefore able to support the translation of *gfpuv* mRNA when co-transcribed with *ble*.

Quantification of GFP suggested that expression using IG4 is greater than IG5, which in turn is greater than IG2, which had very low expression (Figure 4.15).

GFP fluorescence was measured using a luminescence spectrometer and plotted in the same way as Figure 4.10 in section 4.2.2 to remove the effect of the peak at 455nm (Figure 4.16). The expression of GFP after IG4 is greater than IG5 and IG2 and matches the results seen in the Western blot analysis.

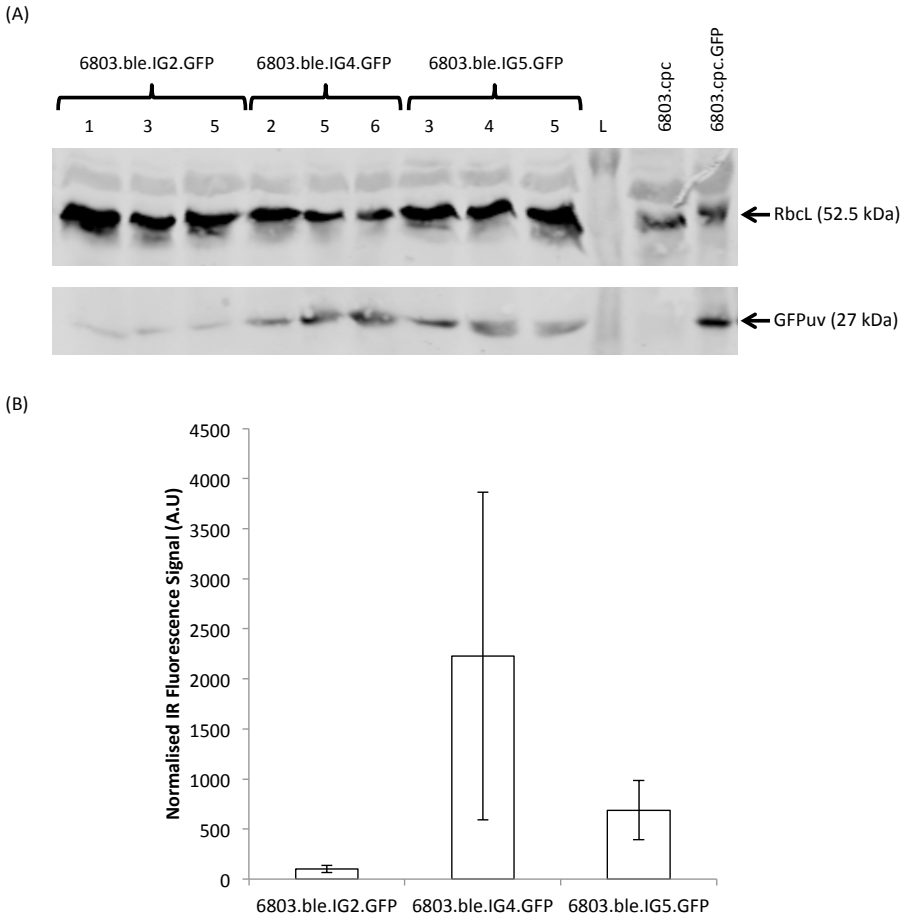


Figure 4.15 Western blot analysis confirms and compares GFPuv expression after IG2, IG4 and IG5 in *Synechocystis* strains expressing the two-gene operons.

(A) Western blot analysis was performed with whole cell extracts with anti-RbcL and anti-GFP antibodies using the Odyssey® Infrared Imaging System for detection. The samples were run on a 12% SDS-PAGE gel. The expected band size for GFP is 27 kDa and 52.5 kDa for RbcL, which acted as loading control. Protein size was determined using the PageRuler™ prestained protein ladder. Three biological replicates of 6803.ble.IG2.GFP, 6803.ble.IG4.GFP and 6803.ble.IG5.GFP were tested. Strains 6803.cpc. and 6803.cpc.GFP acted as negative and positive controls, respectively. (B) The GFPuv signal was quantified using the Odyssey® Infrared Imaging System. The GFPuv IR fluorescence signal was normalised against the signal of RbcL. The average normalised IR signal was plotted. Error bars are ± one SD.

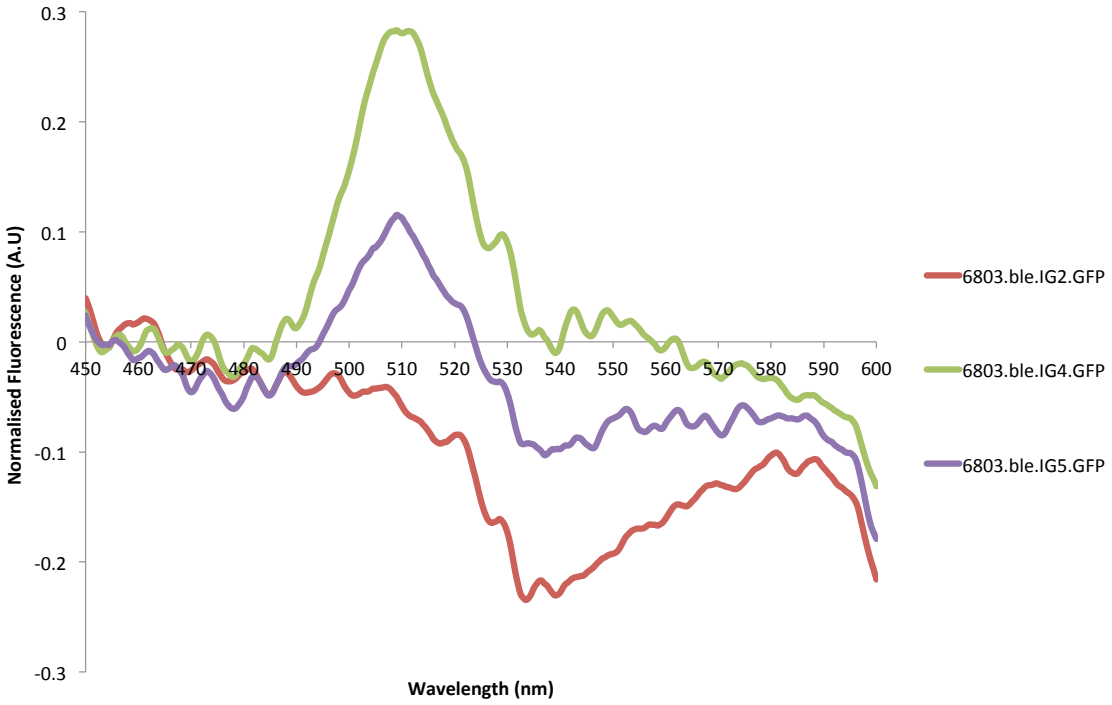


Figure 4.16 GFPuv fluorescence measured after various IGs (IG2, IG4, IG5) under the control of promoter, P_{cpc560} in transformants of *Synechocystis*.

Cultures of similar OD_{730} were excited at 395 nm and the emission spectrum between 450 – 600 nm was plotted. The absorbance of each wavelength was normalised to the absorbance value at 455 nm and the spectrum of the control strain (6803.cpc) was then subtracted from these values. The emission peak at 509 nm is GFPuv.

Interestingly, preliminary data comparing 6803.ble.IG0bp.GFP and 6803.ble.IG1bp.GFP confirmed that gfpuv could be expressed without an IG to give readily detectable levels of the protein (Figure 4.17). Quantification of the GFP does not show a statistically significant difference between a zero base gap and a one base gap between the *ble* and *gfpuv* coding sequences.

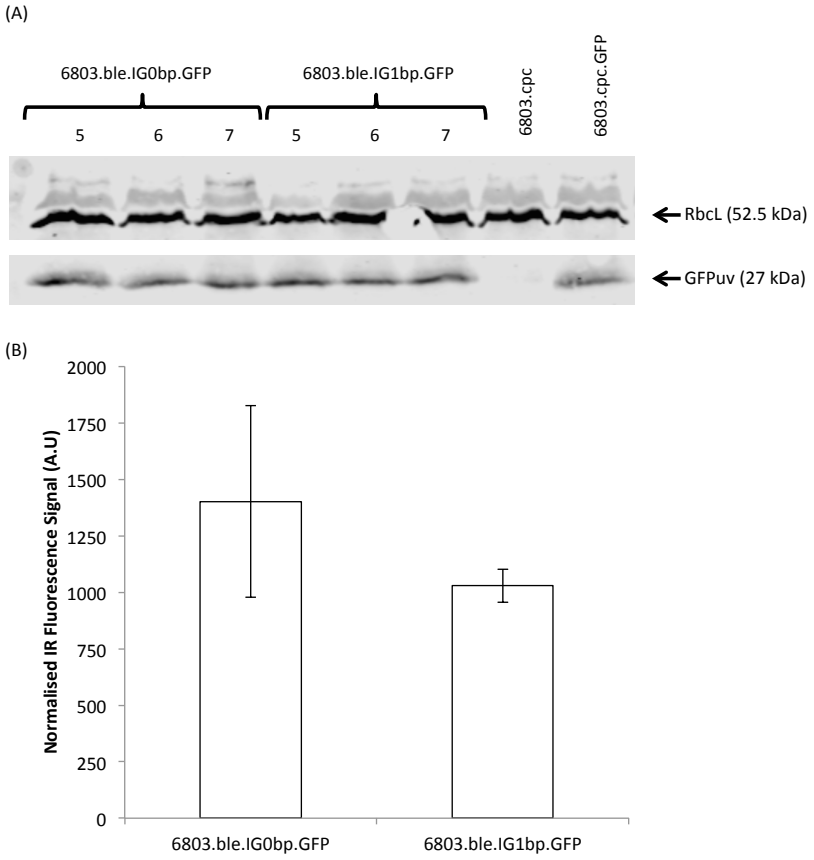


Figure 4.17 Western blot analysis confirms and compares GFPuv expression after IG0bp and IG1bp in the *Synechocystis* expressing the two-gene operons.

(A) Western blot analysis performed with whole cell extracts with anti-RbcL and anti-GFP antibodies using the Odyssey® Infrared Imaging System for detection. The samples were run on a 12% SDS-PAGE gel. The expected band size for GFP is 27 kDa and 52.5 kDa for RbcL, which acted as loading control. Protein size was determined using the PageRuler™ prestained protein ladder. Three biological replicates of 6803.ble.IG0bp.GFP and 6803.ble.IG1bp.GFP were tested. Strains 6803.cpc and 6803.cpc.GFP acted as negative and positive controls, respectively. (B) The GFPuv signal was quantified using the Odyssey® Infrared Imaging System. The GFPuv IR fluorescence signal was normalised against the signal of RbcL. The average normalised IR signal was plotted. Error bars are ± one SD.

When all the IGs are compared again by Western blot analysis, a ranked order of IG4 > {IG0bp, IG1bp, IG5} > IG2 can be produced (Figure 4.18).

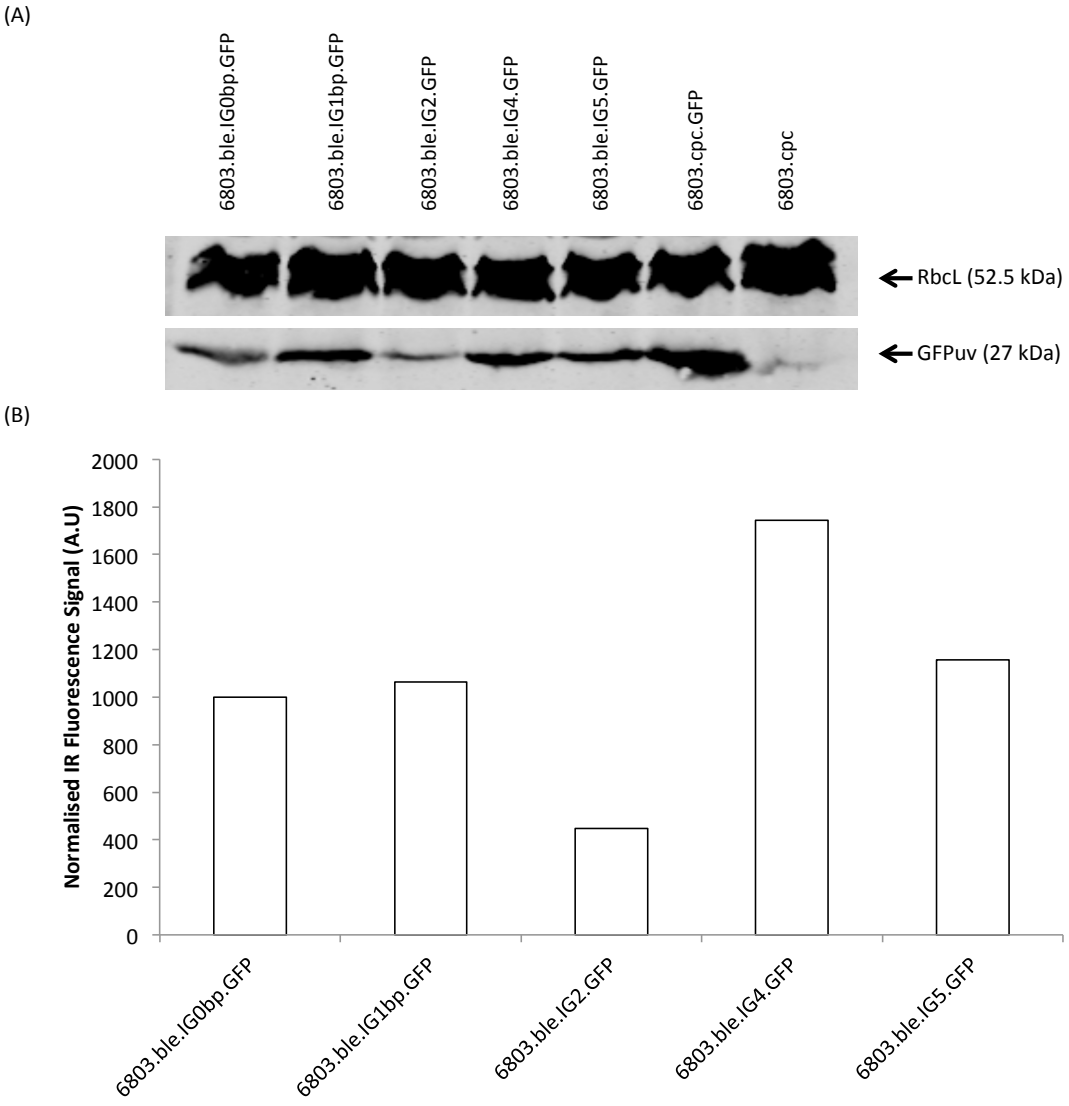


Figure 4.18 Western blot analysis confirms and compares GFPuv expression after IG2, IG4, IG5, IG0bp and IG1bp in the *Synechocystis* expressing the two-gene operon.

(A) Western blot analysis performed with whole cell extracts with anti-RbcL and anti-GFP antibodies using the Odyssey® Infrared Imaging System for detection. The samples were run on a 12% SDS-PAGE gel. The expected band size for GFP is 27 kDa and 52.5 kDa for RbcL, which acted as loading control. Protein size was determined using the PageRuler™ prestained protein ladder. A representative of 6803.ble.IG0bp.GFP, 6803.ble.IG1bp.GFP, 6803.ble.IG2.GFP, 6803.ble.IG4.GFP and 6803.ble.IG5.GFP were tested. Strains 6803.cpc. and 6803.cpc.GFP acted as negative and positive controls, respectively. (B) The GFPuv signal was quantified using the Odyssey® Infrared Imaging System. The GFPuv IR fluorescence signal was normalised against the signal of RbcL and plotted.

4.2.3.4 Further creation of transgenic strains

From the results in the previous section, the expression of *gfpuv* was seen after a 0/1 base space between the *ble* stop codon and the start of *gfpuv*. This begs the question of how

translation of gfpuv is achieved in the absence of a known RBS. Further inspection of the *ble* sequence suggests that there may be a cryptic RBS (Table 4.5) at the end of *ble* that enables the ribosome to bind and initiate translation of *gfpuv*.

Table 4.5 3' end of ble with the potential RBS

The sequence of ble compared to the optimal sequence for the RBS with the 3' end of the splB sequence.

Optimal Complementary SD (5'-3')	_____AAAGGAGGUGAU___
3' end of *ble* (5'-3')	GCGTGCACTTCGTGGCCGAGGAGCAGGACTGA
3' end of *splB* (5'-3')	AATACCCATACGATGTTCCAGATTACGCTTAA

To see whether the putative RBS in *ble* had an effect on the expression of *gfpuv*, the two-gene operon was redesigned so that the gene preceding it did not have a SD-like sequence. From genes that were easily accessible within the lab, *splB*, a *Staphylococcus aureus* gene codon optimised for *Chlamydomonas reinhardtii* that encodes a serine proteasewas chosen (Table 4.5).

Primers used to transcriptionally couple *splB* and *gfpuv* were designed in a similar manner to those used to amplify *ble* when the intergenic region was zero base and one base (for further details see 2.8.2).

WT *Synechocystis* was transformed with pLAH.splB.IG0bp.GFP and pLAH.splB.IG1bp.GFP, to create strains 6803.splB.IG0bp.GFP and 6803.splB.IG1bp.GFP, respectively. After selection for kanamycin resistance, four putative transformants of 6803.splB.IG0bp.GFP and 6803.splB.IG1bp.GFP were checked by PCR. All putative transformants tested contain the splB_IGX_GFP operon at the *psbAII* site and all appear homoplasmic (*i.e.* all copies of the WT genome have been lost from the transformants) (Figure 4.19).

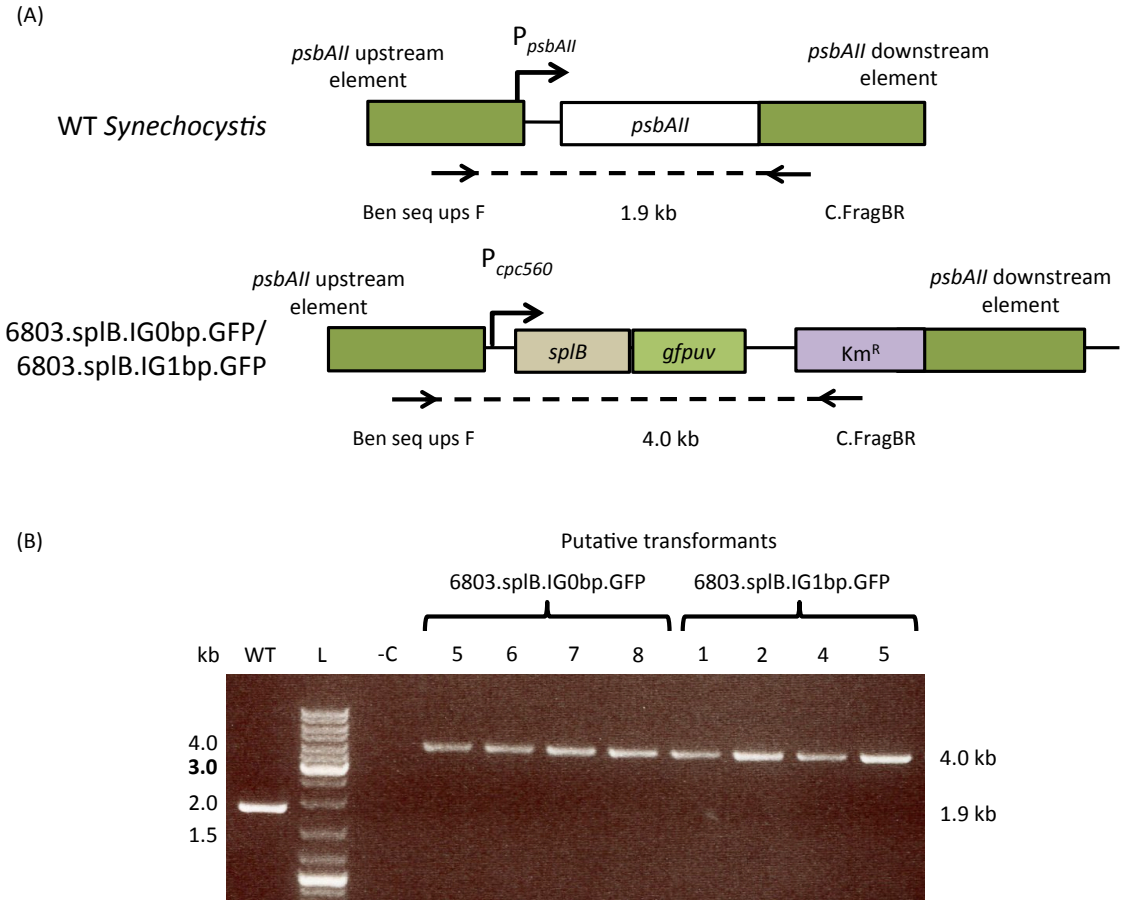


Figure 4.19 PCR screen of putative transformants in *Synechocystis* containing the two-gene operon of *splB* and *gfpuv* with IG0bp/IG1bp under the control of P_{cpc560}

(A) Schematic showing the binding sites and expected fragment sizes for WT *Synechocystis* and 6803.splB.IGX.GFP strains. (B) PCR analysis of putative transformants of 6803.splB.IGX.GFP strains. WT Synechocystis (WT) and no DNA (-C) were used as controls. Transformants were confirmed with a band at 4.0 kb and all appear homoplasmic as they lack any detectable WT band at 1.9 kb.

4.2.3.5 Testing levels of GFP in the new *splB-gfpuv* configurations

To determine GFP fluorescence levels, two transformants of 6803.splB.IG0bp.GFP and three transformants of 6803.splB.IG1bp.GFP were grown along with 6803.cpc (negative control) and 6803.ble_IG4_GFP (positive control). In an attempt to reduce the peak at 455 nm from obscuring the GFP peak at 509 nm, cells were grown in low light (methods) and the cells were harvested and resuspended to the same cell concentration, but were analysed in same manner as (4.2.3.3). The results show translation of *gfpuv* still occurs in the presence of a single base between *ble* and *gfpuv* or no IG region (Figure 4.20). The level of GFP produced appears to be greater when there is a single base in comparison to a zero base gap (Figure 4.20).

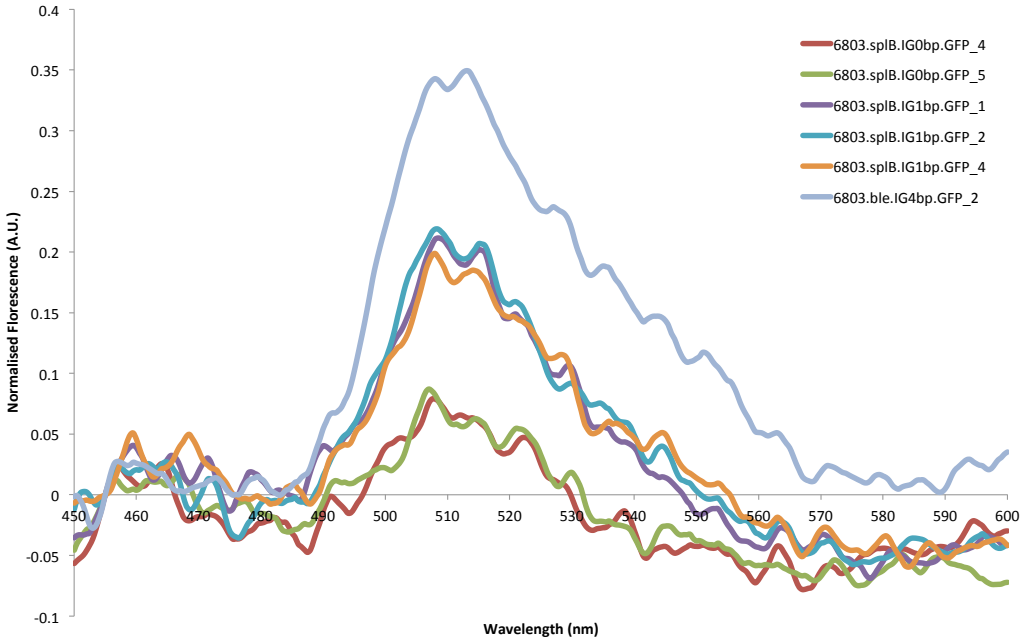


Figure 4.20 GFPuv fluorescence measured after splB and IG0bp or IG1bp under the control of promoter P_{cpc560} in transformants of *Synechocystis*.

The emission peak at 509 nm is from GFPuv. Cultures were prepared to the same cell concentration, excited at 395 nm and the emission spectrum between 450 – 600 nm was plotted. The absorbance of each wavelength was normalised to the absorbance value at 455 nm and the spectrum of the control strain (6803.cpc) was then subtracted from these values. 6803.ble.IG4.GFP acted as the positive control.

To compare all the IGs were grown 6803.cpc (negative control), 6803.cpc.GFP (positive control), 6803.splB.0bp.GFP, 6803.splB.1bp.GFP, 6803.ble.IG0bp.GFP, 6803.ble.IG1bp.GFP, 6803.ble.IG2.GFP, 6803.ble.IG4.GFP and 6803.ble.IG5.GFP. The cultures cells were prepared as described when testing 6803.splB.IG0bp.GFP and 6803.splB.IG1bp.GFP (in low light). The positive control, 6803.cpc.GFP, gave the highest level of expression, followed by IG4, 1bp (ble), IG5, 0bp (ble), 1bp (splB), 0bp (splB) then IG2 (Figure 4.21). The expression of *GFPuv* after *ble* when there is a 0/1 bp gap is greater that after *splB*. This suggests the sequence preceeding *gfpuv* does have an effect on its expression when there is no IG present.

Although the results have not been validated by Western blot analysis, IG4 consistently gives the highest level of GFP, whilst IG5 and IG2 gave very low levels. The expression of *gfpuv* after IG5, IG0bp and IG 1bp (after *ble*) are similar.

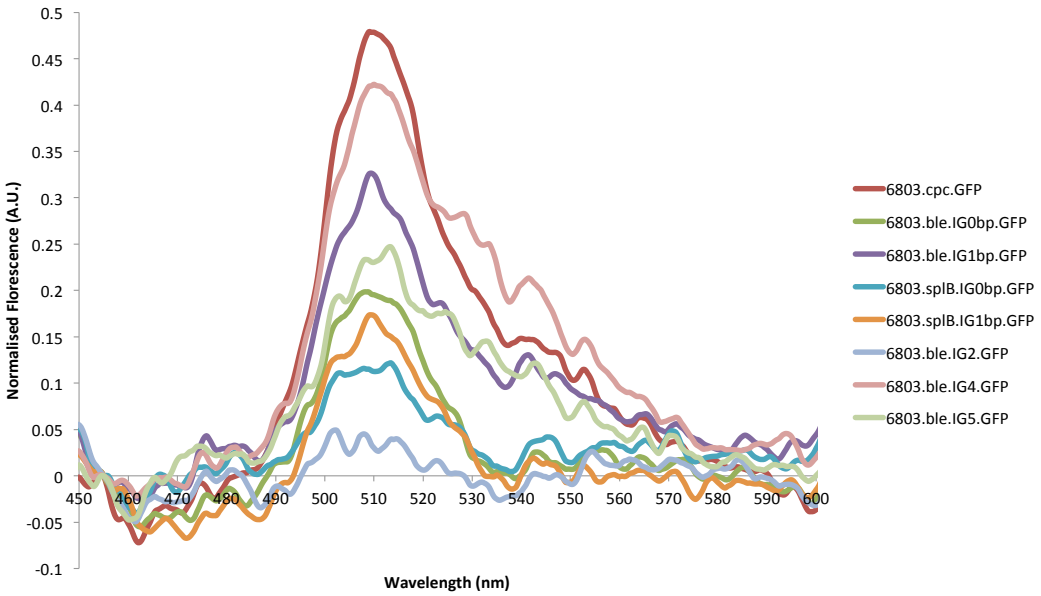


Figure 4.21 GFPuv fluorescence measured in all strains under the control of promoter P_{cpc560} in transformants of *Synechocystis*.

The emission peak at 509 nm is from GFPuv. Cultures were prepared to the same cell concentration and excited at 395 nm and the emission spectrum between 450 – 600 nm was plotted. The absorbance of each wavelength was normalised to the absorbance value at 455 nm and spectrum of the control strain (6803.cpc) was then subtracted from these values. 6803.cpc was used as positive control.

4.3 Discussion and future work

4.3.1 Transcriptional controls

4.3.1.1 Nickel inducible promoter

The metabolic engineering of a microorganism such as *Synechocystis* requires the coordinated expression of multiple transgenes. Furthermore, that expression should be tightly controllable since the synthesis of novel metabolites can typically place a metabolic burden on the cell, and in the worst-case scenarios prove toxic. The nickel-inducible promoter, P_{nrsB}, represents one possible tool for control of multiple transgenes assembled as an operon. The results presented here show that the expression of the reporter genes *ble* and *aadA* can be controlled by Ni^{2+} concentration. However, the work also highlighted two potential issues. Firstly, under the low Ni^{2+} conditions used, the promoter appears not to be completely 'off' as slow growth on spectinomycin-containing medium was observed after nine days. This could also be a result of weak expression from a separate cryptic promoter upstream of the *aadA* gene. From the literature the *nrsBACD* promoter is thought to have two promoters, one is induced by presence of Ni^{2+}, and the other is constitutive, found further upstream, and is present in P_{nrsB} in the pLAH.nrsB construct, that produces very low levels of transcripts (López-Maury et al., 2002). After nine days, the low levels of transcript may express enough AadA to grow in the presence of spectinomycin.

The second issue is certainly the result of a cryptic promoter, with the Ni^{2+}–regulated expression of the third gene in the operon, *aphA-6*, being 'over-ridden' by expression from another promoter somewhere upstream of *aphA-6*, but downstream of the translational start of *aadA*. Two potential cryptic promoters were found upstream of *aphA-6*, in *aadA* and IG2, and one or both are likely responsible for the transcription of *aphA-6* in the absence of added Ni^{2+}. Although transcription in *E. coli* differs to *Synechocystis*, BPROM, the software used to predict the cryptic promoter, searched for σ^{70} promoters, which are also found in *Synechocystis* (Imamura and Asayama, 2009). One of the potential promoter regions covers part of the *aadA* gene and IG2. To determine whether this potential promoter that is responsible for the expression *aphA-6*, the intergenic region could be swapped for another and the expression of *aphA-6* could be again tested in the presence of kanamycin with and without nickel.

4.3.1.2 Comparison of the *cpc560* promoter to the commonly used *psbAII* promoter

The expression of the GFP gene was found to be approximately 3-fold greater under P_{cpc560} than P_{psbAII}, making this new promoter element a useful DNA part in the cyanobacterial 'toolkit'. However, expression under P_{cpc560} was less than expected considering expression of two other genes, as reported by Zhou et al. (2014), resulted in the recombinant protein being approximately 15% of the total soluble protein and clearly visible on an stained SDS-polyacrylamide gel.

The construction of the strains presented in this chapter and the difference in growth conditions may have contributed to the modest levels of expression seen with P_{cpc560}. In the work described by Zhou et al. (2014), the *rbc* terminator sequence was used after the two test genes (*Ter* and *DldhE*), and the genes were also integrated at the *pta* locus that encodes phosphotransacetylase rather than the *psbAII* locus. The *pta* gene is involved in the acetate synthetic pathway and knocking out *pta* does not affect growth (Zhou et al., 2014). These differences along with differing growth conditions could all have an effect on transcription levels. To make a true comparison the method used by Zhou et al. would have to be attempted. In particular, it would be interesting to see the effect of higher light levels on each promoter, as transcription from both promoters is induced by light (Gill et al., 2002; Lindberg et al., 2010).

4.3.2 Translational control within an operon

The expression of GFP was detected for all the IGs tested in the two-gene operon under the control of P_{cpc560}. The IG that gave the highest expression was IG4, which was designed to contain a RBS site that was complementary to the 16S rRNA of *Synechocystis*. Expression was also found to be possible even when there was no bases or a single base between the stop codon of the first gene and the start codon of the second gene, suggesting that a process of translational coupling might be operating in which the ribosome is able to associate with the start codon of the second gene once it reaches the stop codon of the first gene.

IG4 has the best-matched SD sequence out of the intergenic regions tested and as such was expected to give the highest level of GFP expression. In prokaryotes, highly expressed genes are more likely to have a strong SD sequence (Ma et al., 2002). The next best IG region was IG5. This IG region is found between a hypothetical protein (*sll0939*) and aspartate transaminase (*sll0938*). IG2 gave the lowest level of GFP expression but was selected because IG2 had been previously shown to work. IG2 is found between *sll1032*

(*ccmN)* and *sll1031* (*ccmM*) part of the gene cluster *ccmKLMN*, which encodes genes that assemble the shell of the carboxysome (Cot et al., 2008).

An interesting observation from the work is the finding that omission of an IG between two coding sequences still results in translation of the downstream sequence. In the case of the *ble-gfpuv* pairing the translation of *gfpuv* could be accounted for by the finding of a cryptic SD sequence at end of *ble*. However, translation of *gfpuv* is still seen when *ble* was replaced by *splB*, which was chosen specifically as it did not contain an obvious SD sequence. Significantly, only 26% of all genes in the *Synechocystis* genome contain the core SD sequence, whereas in *E. coli* this figure is 57% (Ma et al., 2002). Abe et al looked briefly at the SD sequence of genes involved in photosynthesis, and found *psbA* and *cpcB* both have SD-like sequences in their 5' UTRs, AGGA and AGGAG, respectively (Abe et al., 2014). This suggests that many genes in *Synechocystis* do not need an SD sequence for translation, although expression may be more effective for those genes that do.

Ideally the work initiated here would be extended to more IGs. These could be quantified in terms of GFP fluorescence levels achieved using the strategy developed to provide a list of IGs that can be inserted between genes when designing a transgenic operon, so that translation of each gene can be balanced to optimise the yield.

Chapter 5 Developing a new method for the production of marker-less transformants in *Synechocystis*.

5.1 Introduction

5.1.1 Marker-less transformants

Selectable markers, often bacterial antibiotic resistance genes, are used to select for rare transformants amongst a large population of treated cells. However, once such transformant lines have been recovered, the selectable marker serves no further purpose and its presence within the transformant genome becomes superfluous. A selection system that results in marker-less transformants is valuable since it avoids biosafety concerns relating to the potential horizontal gene transfer of these resistance genes to other bacteria (Tan et al., 2013). In addition, there is a reduced 'metabolic load' on the transformant line as it is not required to replicate and express an unnecessary gene. Finally, multiple cycles of transformation of a strain are not limited by the number of selectable markers available (Cheah et al., 2013; Viola et al., 2014).

The current methods of producing marker-less transformants in different strains of cyanobacteria are discussed below.

5.1.2 Production of marker-less transformants using the expression of *sacB*

The *sacB* gene from *Bacillus subtilis* encodes levansucrase, a secreted enzyme. The expression of *sacB* is regulated by a *cis*-acting regulatory locus, *sacR*, and is induced in the presence of sucrose (Steinmetz et al., 1985). Levansucrase transfers fructosyl residues from sucrose via hydrolysis to different acceptors such as levans; other acceptors include water and other sugars (Steinmetz et al., 1985). Levans are branched polymers made up of fructosyl residues that are added to further by levansucrase (Steinmetz et al., 1985). The expression of *sacB* in the presence of sucrose in Gram-negative bacteria such as *E. coli* is lethal to the cell. The mechanism of toxicity is not clear, it might be due to the transfer of the fructosyl residue to metabolically important acceptors or it might be due to the accumulation of levans in the periplasm that cannot be cleaved or exported out of the cell (Gay et al., 1983). As the peripheral cell structure of Gram-negative bacteria is similar to cyanobacteria, the *sacB*-sucrose method was established in *Anabaena sp.* Strain PCC 7120 and *Synechocystis* (Cai and Wolk, 1990; Lagarde et al., 2000).

The *sacB* – sucrose method is one of the more commonly used methods for producing marker-less transformants in *Synechocystis* and is described in Figure 5.1. Usually the method requires two transformation steps and two transforming plasmids, however a method has recently been developed where only a single transforming plasmid is required (Viola et al., 2014).

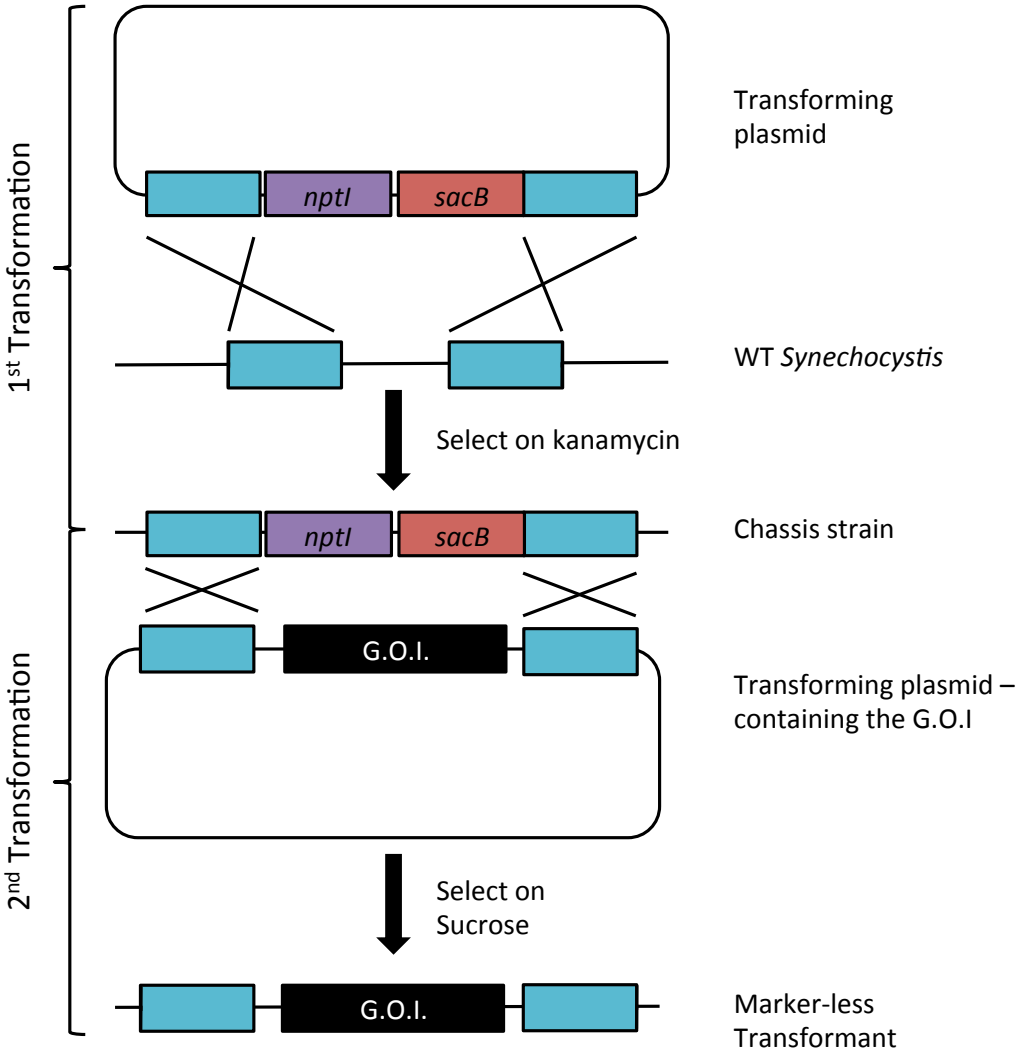


Figure 5.1 Pipeline for the production of marker-less transformants using the *sacB*–sucrose method.

In the first transformation step, *sacB* alongside an antibiotic resistance gene (in this case *nptI*, a gene conferring resistance to kanamycin) is integrated by homologous recombination into the genome of *Synechocystis*. The creation of the chassis strain containing *sacB* is selected for on kanamycin. The gene-of-interest, G.O.I, replaces the *sacB*/*nptI* gene cassette in the second transformation, with the marker-less transformants selected on sucrose for the loss of *sacB*.

5.1.3 Other methods used in cyanobacteria to produce maker-less transformants

Although a number of methods have been developed for the production of marker-less transformants (Table 5.1), there are three distinct strategies used. One strategy uses counter-selection and is used by the *sacB* – sucrose method. Initially an antibiotic resistance gene, e.g. *nptI*, is introduced into the genome alongside a conditionally toxic

gene, e.g. *sacB*. The toxic gene is then expressed, so only the transformants that no longer contain the toxic gene and the antibiotic resistance gene are able to survive.

Table 5.1 Overview of the various methods used to produce marker-less transformants in cyanobacteria.

Adapted from (Branco dos Santos et al., 2014). Counter selection systems (in blue), site-specific recombinase systems (in green) and a mutant-complementation-based system (in white) have been used to produce marker-less transformants in various strains of cyanobacteria.

Method	Strain tested	Reference
sacB - sucrose	*Synechocystis* sp. PCC 6803 *Anabaena* sp. strain PCC 7120	(Lagarde et al., 2000) (Cai and Wolk, 1990)
sacB - sucrose (one-step)	*Synechocystis* sp. PCC 6803	(Viola et al., 2014)
mazF - nickel	*Synechocystis* sp. PCC 6803	(Cheah et al., 2013)
acsA - acrylate	*Synechococcus* sp. PCC 7002	(Begemann et al., 2013)
upp – 5-fluorouracil	*Synechococcus elongatus* PCC 7942 *Synechocystis* sp. PCC 6803 *Synechococcus* sp. PCC 7002	(Aikens and Turner, 2013) (Aikens and Turner, 2013) (Xu and Green, 2012)
rps12 - streptomycin	*Synechococcus elongatus* PCC 7942	(Matsuoka et al., 2001)
Cre/ loxP	*Anabaena* sp. PCC 7120	(Zhang et al., 2007)
Flp/ FRT	*Synechocystis* sp. PCC6803 *Synechococcus elongatus* PCC 7942	(Tan et al., 2013) (Tan et al., 2013)

The *mazF* – nickel method uses the same strategy and was developed in *Synechocystis* to be used in other cyanobacteria strains that can integrate DNA by homologous recombination but are unable to use the *sacB* – sucrose method, as they cannot uptake sucrose (Cheah et al., 2013). The *E. coli mazF* gene encodes an endoribonuclease that cleaves mRNA at the ACA triplet sequence thereby inhibiting all protein synthesis (Cheah et al., 2013). Expression of *mazF* under the control of a nickel inducible promoter acts as

the conditionally lethal gene, therefore marker-less transformants are selected on nickel for the loss of *mazF* (Cheah et al., 2013).

Acrylate inhibits *Synechococcus* sp. PCC 7002, however, the loss of function of *acsA,* which encodes acetyl-CoA ligase, was shown to overcome this inhibition (Begemann et al., 2013). Marker-less transformants can therefore be produced by selecting for the loss of *acsA*, introduced alongside an antibiotic resistance cassette, in the presence of acrylate. This method uses a similar strategy to the *sacB* – sucrose method; however, the original *acsA* gene needs to be knocked out before the strain can be used to produce marker-less transformants. Similarly, the *upp* – 5-fluorouracil method, used in *Synechocystis*, *Synechococcus* 7942 and *Synechococcus* 7002, has the same issue and requires the deletion of *upp* (Aikens and Turner, 2013; Xu and Green, 2012). The *upp* gene encodes uracil phosphoribosyltransferase, an enzyme that usually converts uracil into UMP but can also accept 5-fluorouracil to produce 5-fluoro-UMP which results in the inhibition of protein synthesis and leads to cell death.

Another strategy uses a strain of *Synechococcus* sp. PCC 7942 that is resistant to streptomycin due to a single mutation in a ribosomal S12 protein encoded by *rps12* (Matsuoka et al., 2001). The reintroduction of WT *rps12* alongside an antibiotic resistance gene makes the strain sensitive to streptomycin. Marker-less transformants are then selected on streptomycin (Matsuoka et al., 2001). Unlike the *sacB* – sucrose method, there are two rounds of positive selection, which is advantageous, however the resulting strain has a mutation in a ribosomal protein and this might compromise translational efficiency and therefore the 'fitness' of the engineered strain.

The third strategy uses site-specific recombinases such as Cre from the *E. coli* bacteriophage P1 and FLP from *S. cerevisiae.* Site-specific recombinases catalyse recombination between two site-specific recombinase recognition sites. Cre recognises the locus of crossover in phage P1 (*loxP*) and FLP recognises FLP recognition target (FRT). Flanking a selectable marker with the specific recognition sites, and expressing the site-specific recombinases, removes the selectable marker (Tan et al., 2013). Cre-*loxP* has been used in *Anabaena* 7120 and the FLP-FRT method has been developed for *Synechocystis* and *Synechococcus* 7942 (Zhang et al., 2007; Tan et al., 2013). Using this method does have some disadvantages: a scar sequence remains and multiple uses of this strategy can result in unexpected and undesirable crossover events.

5.1.4 Could the expression of *codA* be used to produce marker-less transformants in *Synechocystis*?

Cytosine deaminase is an enzyme involved in converting cytosine and water into uracil and ammonia (Figure 5.2) (Danielsen et al., 1992). Cytosine deaminase is found in prokaryotes, fungi and some eukaryotic microorganisms but is not present in higher plants (Danielsen et al., 1992; Serino and Maliga, 1997). In some cases cytosine deaminase can also accept 5-fluorocytosine (5-FC) as a substrate to produce 5-fluorouracil, a compound that leads to cell death by inhibition of DNA and protein synthesis (Figure 5.2)(Vermes et al., 2000). This has led to the use of 5-FC as a negative-selectable marker in a number of organisms (Stougaard, 1993; Kobayashi et al., 1995; Serino and Maliga, 1997; Dubeau et al., 2009; Young and Purton, 2014).

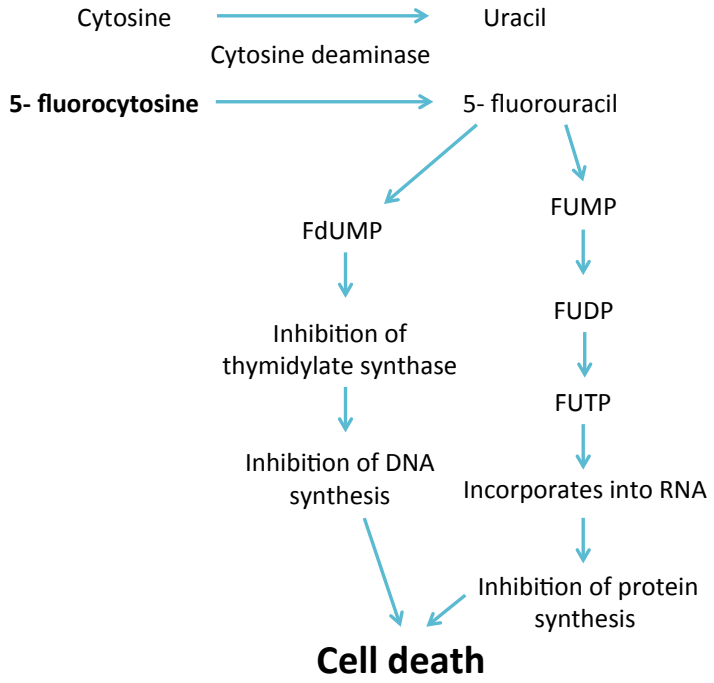


Figure 5.2 Pathway and mode of action of 5-fluorocytosine

Adapted from (Vermes et al., 2000). Abbreviations: 5-FC, 5-fluorocytosine; 5-FU, 5-fluorouracil; FUMP, 5-fluorouridine monophosphate; FUDP, 5-fluorouridine diphosphate; FUTP, 5-fluorouridine triphosphate; FdUMP, 5-fluorodeoxyuridine monophosphate.

Since *codA* and 5-FC can be used as a method of negative selection, it could potentially be used to produce marker-less transformants in *Synechocystis* in a similar method to the *sacB* – sucrose method.

5.1.5 Aims and objectives

The initial objective was to introduce *sacB* at a neutral site in the genome of *Synechocystis* in order to produce a chassis strain that could be used to introduce genes and operons without the presence of an antibiotic resistance marker using the *sacB* - sucrose method. Attempts to use the *sacB* – sucrose method were initially unsuccessful so an alternative method using the expression of a non-native *codA* gene in the presence of 5-FC was developed. The work presented in this chapter describes the attempts of using the *sacB* – sucrose method and the *codA*-5-FC method for the production of marker-less transformants in *Synechocystis*.

5.2 Results and discussion

5.2.1 Creation of chassis strains containing *sacB* gene

To produce marker-less transformants at a neutral site in *Synechocystis* using the *sacB* – sucrose method, it was decided to use the strategy presented in Figure 5.1. Once a neutral site was selected, the various expression plasmids would be created. Initially the *cat* gene encoding chloramphenicol resistance was chosen as the G.O.I.

The strategy used to construct each expression plasmid is shown in Figure 5.3.

5.2.1.1 Identification of a neutral site

To create a standard chassis strain for metabolic engineering it is necessary to identify a region of the genome where transgenes can be targeted without affecting the expression of any native genes. CyanoBase was used to search the genome of *Synechocystis* sp. PCC 6803 for potential neutral sites. Non-coding regions greater than 400 bp between genes that are oriented in a divergent or convergent orientation, and which also contain a unique restriction enzyme site suitable for cloning (that is also absent from the pJET1.2/blunt cloning vector) approximately half way into the non-coding region, were shortlisted as potential neutral sites.

The 705 bp non-coding region between *slr1340* and *sll1255* (genes organised in a convergent orientation), containing the blunt-ended HpaI and sticky-ended NheI restriction sites, was chosen as the neutral site. According to CyanoBase, the gene *slr1340*, encodes an unknown protein, and *sll1255*, encodes a putative transposase.

5.2.1.2 Creation of pIR706

Plasmid pIR706, containing a clone of the neutral site, was created to introduce DNA by homologous recombination at this site. Gradient PCR using IR706F and IR706R (annealing temperature 58, 60, 68°C) was used to amplify the chosen neutral site and part of the *slr1340* and *sll1255* coding region to ensure the region of homology either side of the NheI restriction site was approximately 500 bp. The fragment was cloned into pJET1.2/blunt cloning vector to produce pIR706. The creation of pIR706 was verified by PCR performed with IR706F and IR706R and by sequencing.

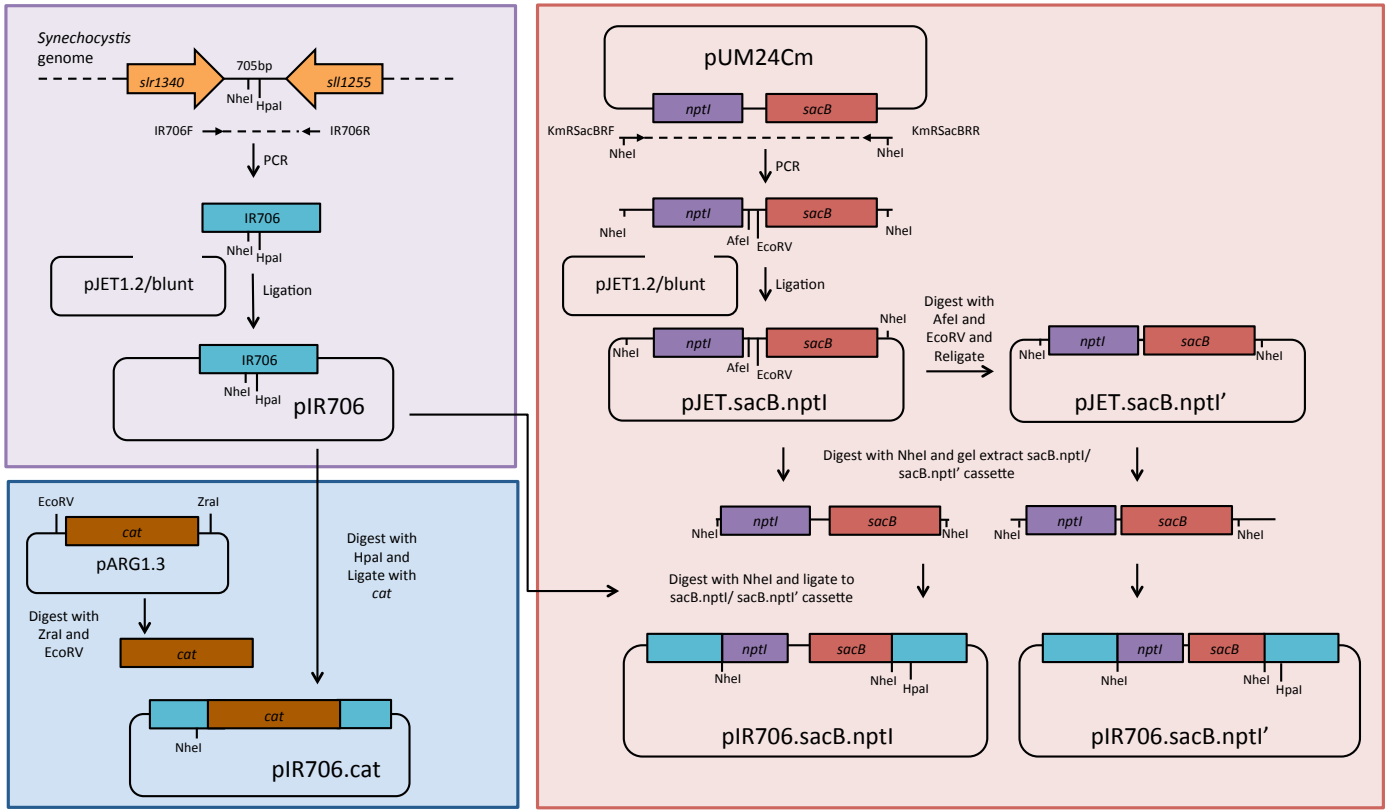


Figure 5.3 Schematic of the plasmids constructed for the production of marker-less transformants in *Synechocystis* using the *sacB* – sucrose method at a neutral site.

The neutral site was cloned from the genome of *Synechocystis* and used to construct pIR706 (in the purple box). This expression vector provided the regions of homology required to introduce *sacB* (in the red box), via pIR706.sacB.nptI and pIR706.sacB.nptI', and *cat* (in the blue box), via pIR706.cat, into the genome of *Synechocystis*.

5.2.1.3 Creation of pIR706.cat

In order to create the test plasmid carrying the *cat* G.O.I., the expression cassette containing the *cat* marker was isolated from pARG1.3 (Debuchy et al., 1989) by digestion with EcoRV and ZraI, and cloned into pIR706 at the HpaI site. PCR with IR706F and IR706R and sequencing confirmed the creation of pIR706.cat.

5.2.1.4 Creation of pIR706.sacB.nptI and pIR706.sacB.nptI'

The pIR706.sacB.nptI and pIR706.sacB.nptI' plasmids were created in order to introduce *sacB,* along with *sacR*, alongside *nptI*, a kanamycin resistance gene, into *Synechocystis* to produce the chassis strain (Figure 5.3). In pIR706.sacB.nptI' a section of DNA between *sacB* and *nptI* was removed. This section of DNA, a possible transposon that is likely to be an artefact from cloning, remains present in pIR706.sacB.nptI.

The *sacB/nptI* gene cassette used to create pIR706.sacB.nptI and pIR706.sacB.nptI' was amplified from pUM24Cm (Ried and Collmer, 1987) via PCR using primers KmRSacBRF and KmRSacBRR. Amplification with these primers ensured that the PCR product was flanked with NheI restriction sites. The purified PCR product was cloned into the pJET1.2/blunt vector. To create pIR706.sacB.nptI, the *sacB/nptI* cassette was digested from pJET.sacB.nptI using NheI and cloned into the NheI site in pIR706. To create pIR706.sacB.nptI', the same steps as above were undertaken, however AfeI and EcoRV were used to remove 457 bp of redundant sequence between *sacB* and *nptI* from the pJET.sacB.nptI cloning vector, prior to insertion into pIR706. Sequencing analysis confirmed the construction of both constructs.

5.2.1.5 Creation of the *sacB* chassis strain for the production of marker-less transformants

WT *Synechocystis* was transformed with pIR706.sacB.nptI and pIR706.sacB.nptI'. Transformant colonies for both plasmids were seen after selection for kanamycin resistance. Eight putative transformants were picked from each transformation and restreaked onto BG-11 medium supplemented with 200 μg/ml kanamycin. Colonies were restreaked four times to ensure all copies of the genome contained the *sacB/nptI* cassette. PCR performed using primers IR706F and IR706R on putative transformants transformed with pIR706.sacB.nptI or pIR706.sacB.nptI' confirmed the integration of the *sacB*/*nptI* cassette (Figure 5.4) The PCR also confirmed that the kanamycin resistant lines were homoplasmic.

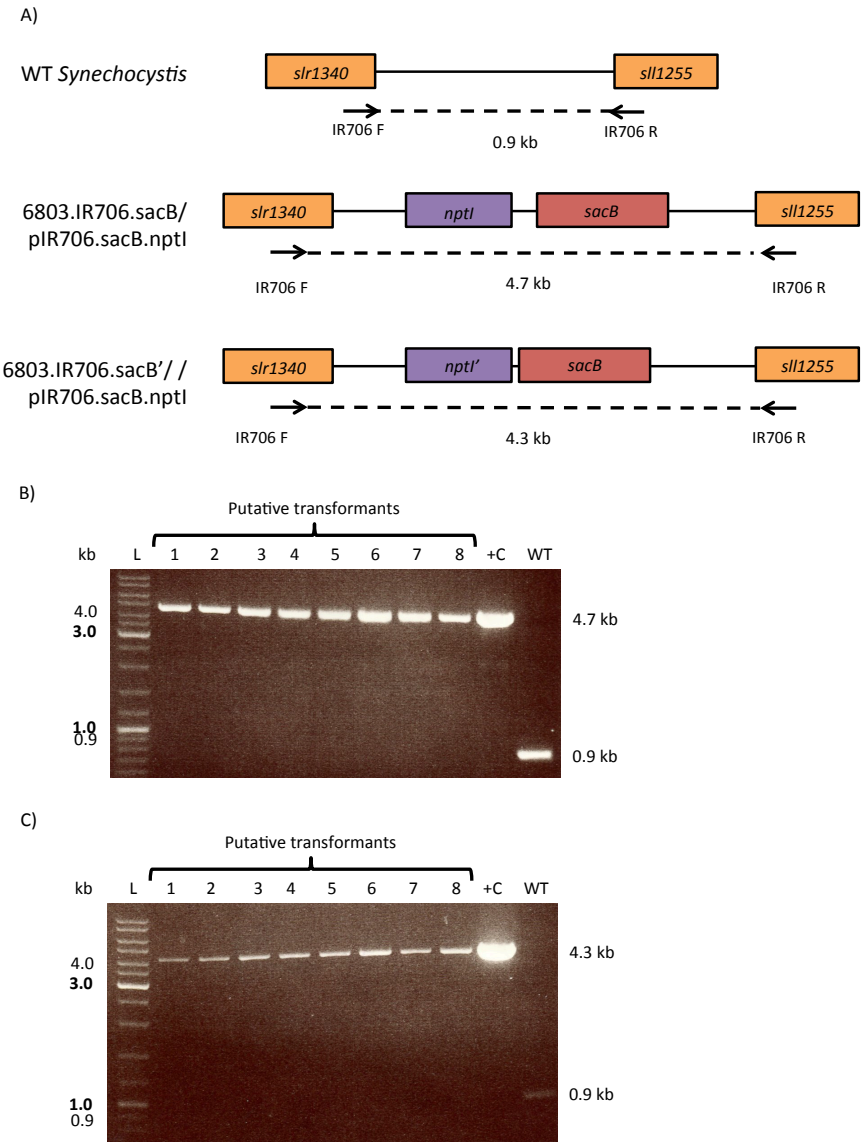


Figure 5.4 PCR screening for successful transformants in *Synechocystis* containing *sacB*.

(A) Schematic showing the binding sites and expected fragment sizes for WT *Synechocystis* and successful 6803.IR706.sacB and 6803.IR706.sacB' strains. (B) Colony PCR screening of putative transformants of 6803.IR706.sacB. WT *Synechocystis* (WT) and a plasmid preparation of pIR706.nptI.sacB (+C) were used as controls. Homoplasmic strains containing *sacB* confirmed by the single band at 4.7 kb. (C) Colony PCR screening of putative transformants of 6803.IR706.sacB'. WT *Synechocystis* (WT) and a plasmid preparation of pIR706.nptI.sacB' (+C) were used as controls. Homoplasmic strains containing *sacB* confirmed by the single band at 4.3 kb.

5.2.1.6 Testing the expression of *sacB*

Growth 'spot' tests were performed to test the sensitivity of strains containing *sacB* towards sucrose. Initial spot tests were performed with three isolated transformants for each strain 6803.IR706.sacB (#1, 2, 7) and 6803.IR706.sacB' (#2, 4, 6) on BG-11, BG-11 supplemented with 200 µg/ml kanamycin and BG-11 supplemented with 5 % (w/v)

sucrose. WT Synechocystis and strain 6803.AII.LS were grown alongside as controls. On BG-11 supplemented with sucrose, all 6803.IR706.sacB (#1, 2, 7) and 6803.IR706.sacB' (#2, 4, 6) transformant lines were able to grow, although the growth of 6803.IR706.sacB_7 was slightly inhibited in comparison to the others (Figure 5.5A).

To confirm this unexpected result, the growth 'spot' test was repeated with an additional transformant of 6803.IR706.sacB (#1, 2, 7 and 8) and WT Synechocystis on BG-11 and BG-11 with 5% (w/v) sucrose. Only a single transformant 6803.IR706.sacB_7 was unable to grow in the presence of 5% sucrose (Figure 5.5B),

The PCR results (Figure 5.4) confirm *sacB*/*nptI* integration at the neutral site, however the results of the growth 'spot' tests (Figure 5.5) demonstrate that the expression of *sacB* does not cause cell death in the presence of sucrose in most of the transformants tested. The *sacB* – sucrose method has been used by a number of groups using the same *sacBR*/*nptI* cassette (Lea-Smith et al., 2013) so the result is unexpected.

5.2.1.7 Sequencing analysis of *sacB*

Sequencing analysis was performed to try and explain why 6803.IR706.sacB_7 appeared to be sensitive to sucrose, but the others strains containing *sacB* were not. The *sacB* gene was amplified from the genomic DNA of 6803.IR706.sacB_1 and 6803.IR706.sacB_7 via PCR using NptI_outF and KmRsacBRR. Sequencing both *sacB* genes with 6803_sacB_F, 6803_sacB_R and KmRsacBRR confirmed that no mutations were present in *sacB* or *sacR* in either strain. The sequencing results therefore do not explain why the two transformants produce different results on the growth 'spot' tests and why 6803.IR706.sacB_1 is able to grow on sucrose.

Owing to these results, work on the *sacB* – sucrose method was put on hold to look at alternative methods for the production of marker-less transformants in *Synechocystis* and was revisited in section (5.2.6).

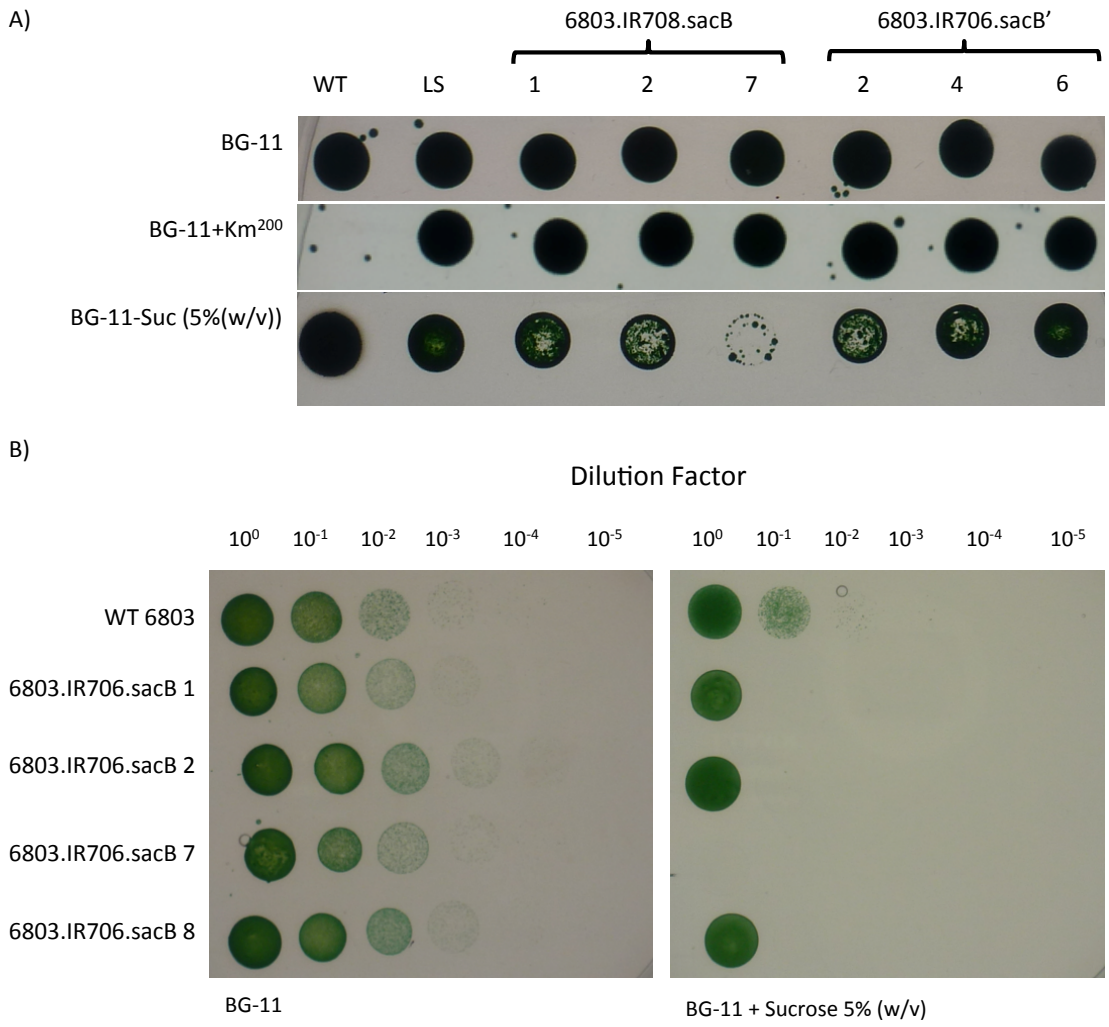


Figure 5.5 Growth 'spot' tests checking sensitivity of *Synechocystis sacB* transformants towards sucrose.

(A) 6803.sacB.nptI and 6803.sacB.nptI' transformants were grown alongside WT *Synechocystis*, (WT) and 6803.AII.LS (LS) which acted as controls, on BG-11, BG-11 supplemented with 200 μg/ml kanamycin and BG-11 supplemented with 5% (w/v) Sucrose. The photograph was taken after 12 days incubation. Kanamycin resistance was seen in all 6803.IR706.sacB and 6803.IR706.sacB' transformants. All 6803.IR706.sacB and 6803.IR706.sacB' transformants were able to grow in presence of sucrose, 6803.IR706.sacB_7 to a lesser degree. (B) Four transformants of 6803.IR706.sacB and wild-type *Synechocystis*, WT 6803, were spotted onto media containing BG-11 and BG-11 supplemented with 5% (w/v) sucrose. A photograph was taken after 4 days incubation. Inhibition of growth by sucrose is only seen in the single transformant 6803.IR706.sacB_7.

5.2.2 Testing the feasibility of a *codA* – 5-FC method for producing marker-less transformants.

Unable to use the *sacB* – sucrose method, an alternative system for producing marker-less transformants was investigated. Previous work in our group demonstrated that a modified cytosine deaminase, encoded by *codA*, could be used as a negative selectable marker in the

chloroplast of *Chlamydomonas reinhardtii* (Young and Purton, 2014). A literature search showed cytosine deaminase had not been used as a negative selectable marker in *Synechocystis*. Therefore, the use of *codA* as a negative selectable marker to produce marker-less transformants in *Synechocystis* was investigated. Figure 5.6 demonstrates how the *codA* – 5-FC method could be transferred to *Synechocystis*.

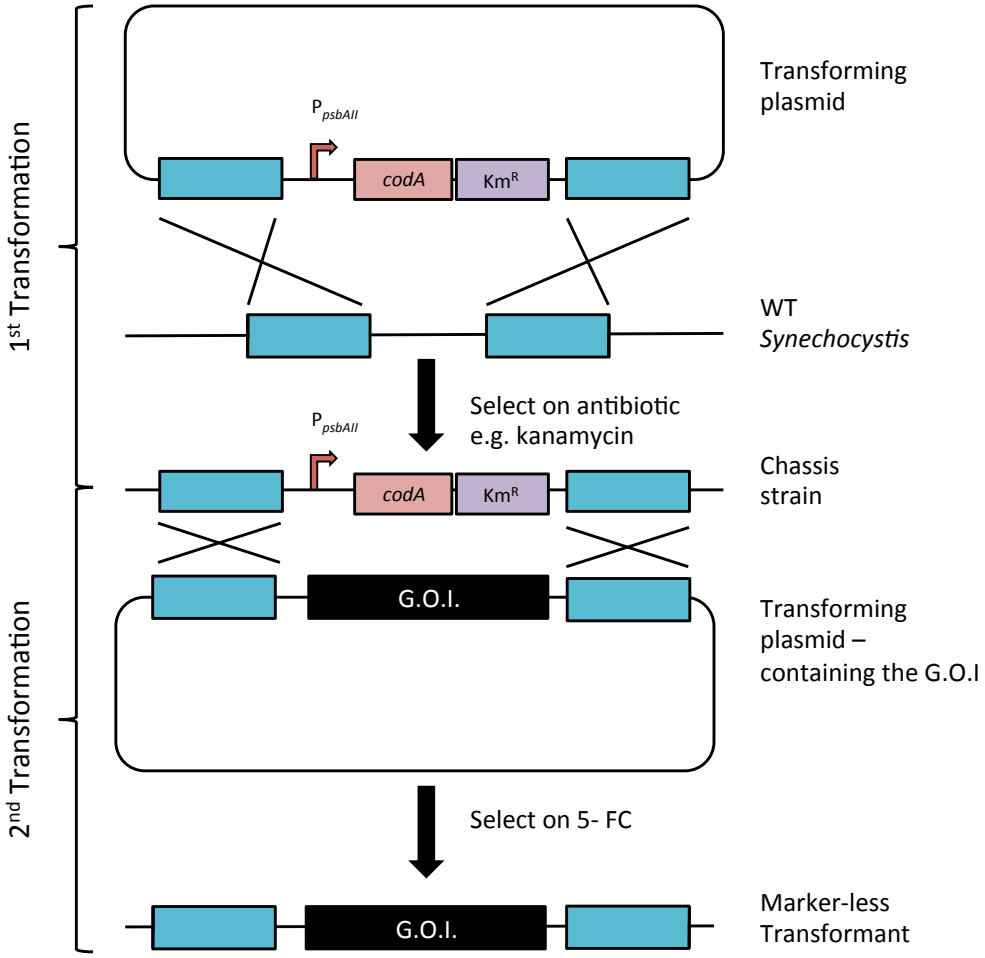


Figure 5.6 Pipeline for the production of marker-less transformants using the expression of *codA* in *Synechocystis*.

In the first transformation step, *codA* alongside a kanamycin resistance gene (*Km*^{R}) is integrated into the genome of *Synechocystis* by homologous recombination. The chassis strain is selected for on kanamycin. The second transformation step selects for marker-less transformants on 5-FC for the loss of the *codA* gene.

5.2.2.1 Endogenous *codA* is not suitable as a negative selectable marker in *Synechocystis*

A search of the *Synechocystis* genome using CyanoBase indicated the presence of a *codA* gene (ID *slr1237*) that is predicted to encode cytosine deaminase and shares 29% amino acid identity with the characterised cytosine deaminase from *E. coli* K12.

An initial growth 'spot' test demonstrated that the growth of WT *Synechocystis* is not inhibited on BG-11 with different concentrations (0.1, 0.5 and 1 mg/ml) of 5-FC (results not shown). Another growth 'spot' test confirmed WT *Synechocystis* could grow in the presence of 2 mg/ml 5-FC (Figure 5.7).

At 2 mg/ml 5-FC the growth of WT *Synechocystis* did appear slower, which suggests that the endogenous enzyme may be active but the substrate specificity towards 5-FC may not be sufficient enough to cause cell death, since the main substrate for the enzyme is cytosine.

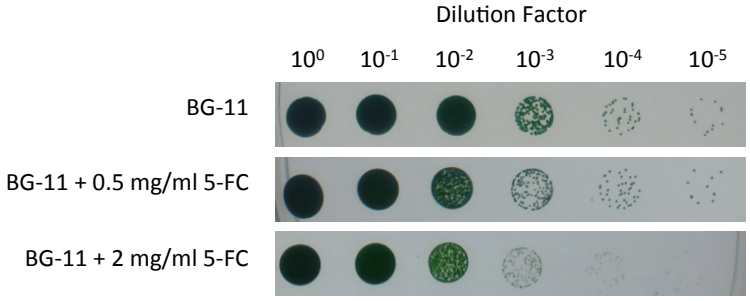


Figure 5.7 Growth 'spot' test testing the sensitivity of WT *Synechocystis* towards 5-FC.

Wild-type *Synechocystis* is able to grow on BG-11 and BG-11 supplemented with 0.5 and 2 mg/ml 5-FC. Photograph taken after 12 days of incubation.

The result demonstrated that the endogenous cytosine deaminase, encoded by *slr1237*, does not confer enough sensitivity to 5-FC to be used to produce marker-less transformants. However, an alternative *codA* encoding a cytosine deaminase that is optimised to accept 5-FC could be used in *Synechocystis*.

5.2.2.2 Expression of a heterologous *codA* gene in *Synechocystis*

The expression of an alternative, synthetic *codA* gene, termed *crCD* (Young and Purton, 2014) was attempted in *Synechocystis.* The design of the synthetic *crCD* gene was based on cytosine deaminase (NP_414871) from *E. coli* strain K12 substrain MG1655. Three amino acid changes (V152A, F316C and D317G) were made to optimise cytosine deaminase to accept 5-FC (Fuchita et al., 2009). A HA-tag was inserted at the C-terminal end and the final gene sequence was codon-optimised for the *Chlamydomonas reinhardtii* chloroplast (Young and Purton, 2014). A BLAST query comparing the amino acid sequence of the native *codA* in *Synechocystis* to *crCD* showed a 28% identity match and a 47% similarity.

The *crCD* gene, referred hereafter as *codA*, was amplified from pRY127d (Young and Purton, 2014) using primers 6803_codA_F and 6803_codA_R. Primers were designed with NdeI (6803_codA_F) and BamHI (6803_codA_R) restriction sites for cloning into the

pLAH.AII expression vector. The *codA* fragment was digested with NdeI and BamHI and inserted into these restriction sites in pLAH.AII . PCR performed with primers Ben.seq.ups.F and Out.GFP.R and sequencing with the same primers confirmed the successful insertion of *codA* to produce pLAH.AII.codA (Figure 5.8).

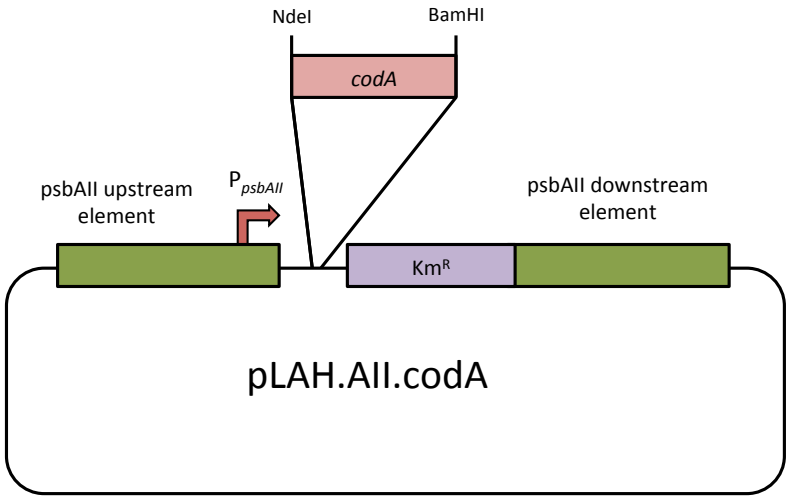


Figure 5.8 Diagram of the expression plasmid pLAH.AII.codA.

The *codA* gene from pRY127d was cloned into the pLAH.AII expression plasmid at the NdeI and BamHI restriction sites to produce pLAH.AII.codA. Cloning in pLAH.AII enables genomic integration of the *codA* gene at the psbAII site under the control of P_{psbAII}.

Synechocystis was transformed with pLAH.AII.codA and transformants appeared after selection for kanamycin resistance. 16 putative transformants were picked and restreaked on BG-11 supplemented with kanamycin (200 µg/ml). A PCR performed on 8 of the 16 putative transformants with Ben.seq.ups.F and C. fragBR confirmed the replacement of *psbAII* with *codA* to produce transformants of 6803.AII.codA (Figure 5.9).

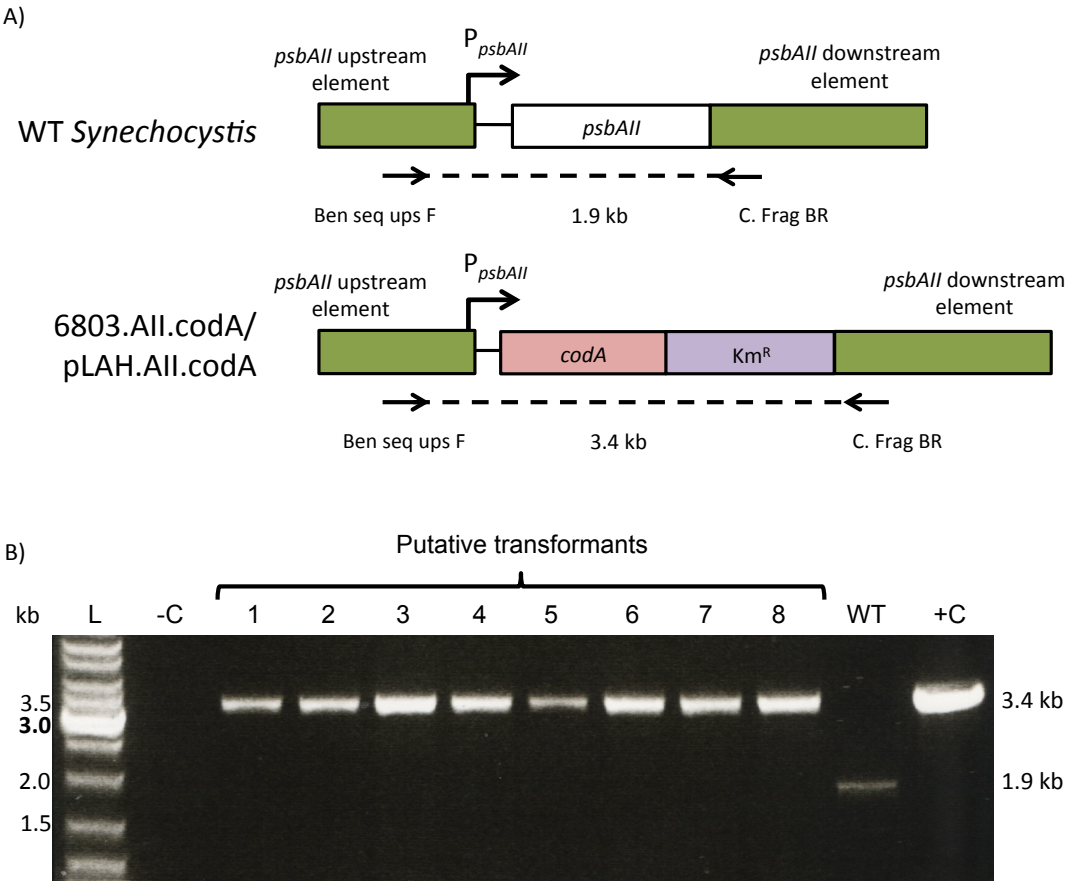


Figure 5.9 PCR screening for successful transformants in the *Synechocystis* containing *codA*.

(A) Schematic showing the binding sites and expected fragments sizes for WT *Synechocystis*, pLAH.AII.codA plasmid and 6803.AII.codA transformants. (B) PCR screening of putative transformants of 6803.AII.codA. WT *Synechocystis* (WT), pLAH.AII.codA, (+C), and no DNA template, (-C), were used as controls. Homoplasmy was confirmed by the lack of a band at 1.9 kb.

5.2.2.3 The heterologous *codA* gene can be used as a negative selectable marker in *Synechocystis*

A growth 'spot' test was performed using a single 6803.AII.codA transformant on BG-11 and BG-11 supplemented with various concentrations of 5-FC (0.5, 1 and 2 mg/ml). The 6803.AII.codA transformant was unable to grow in the presence of 5-FC but was able to grow in the absence of 5-FC (Figure 5.10). WT *Synechocystis* was grown as a control and was able to grow in the presence of 5-FC at 0.5 mg/ml, but growth was restricted at higher concentrations.

As the expression of cytosine deaminase in the presence of 5-FC leads to cell death (Figure 5.10), the results confirm the expression of *codA* under the P_{psbAII} in *Synechocystis*. The

optimised cytosine deaminase can be used as a negative selectable marker in *Synechocystis* and could be used to produce marker-less transformants in a similar method to *sacB.*

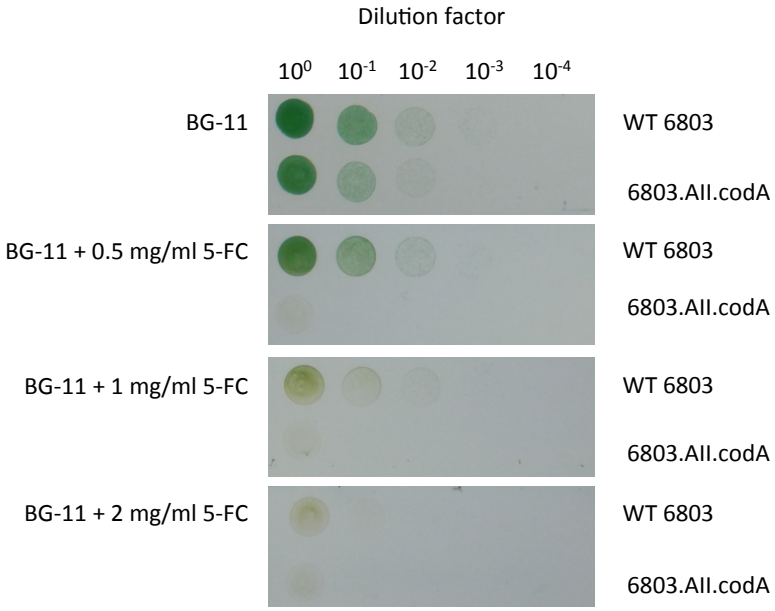


Figure 5.10 Growth spot test checking the expression of *codA* in 6803.AII.codA transformants.

WT *Synechocystis* (WT 6803) and a successful 6803.AII.codA transformant were grown on BG-11 and BG-11 supplemented with 0.5, 1, and 2 mg/ml 5-FC. Photographs were taken after 5 days of incubation. Growth of 6803.AII.codA is inhibited in the presence of 5-FC.

5.2.3 Creation of chassis strain of *Synechocystis* containing *codA* for the production of marker-less transformants

Once the feasibility of the *codA* -5-FC method had been demonstrated, the next step was to create the chassis strain, as shown in Figure 5.6, containing *codA* alongside an antibiotic resistance gene, at the chosen neutral site.

5.2.3.1 Creation of pIR706.codA

The *codA* gene, including the P_{psbAII} promoter was amplified together with the kanamycin resistance cassette, Km^R, from pLAH.AII.codA via PCR using primers pLAH_codAF and pLAH_codAR. Both primers were designed to produce NheI restriction sites at the ends of the PCR product for insertion into pIR706. The PCR product was digested with NheI and cloned into the NheI site in pIR706 to produce pIR706.codA (Figure 5.11). PCR performed with pLAH_codAF and pLAH_codAR, and sequencing with primers IR706F, IR706R and codAmidF confirmed the construction of pIR706.codA.

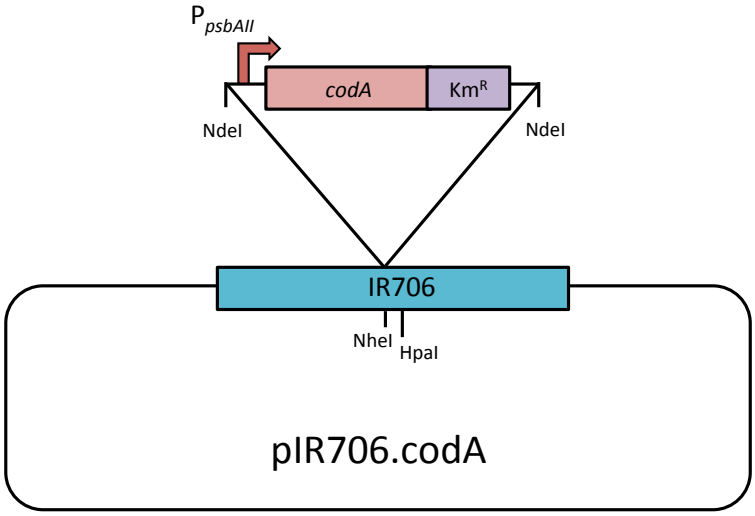


Figure 5.11 Diagram of expression plasmid pIR706.codA.

The modified *codA* gene under the control of the P_{psbAII} was amplified alongside the kanamycin resistance gene and cloned into the NheI site in pIR706 to produce the transforming plasmid pIR706.codA.

5.2.3.2 Expression of *codA* in *Synechocystis* at neutral site

Synechocystis was transformed with pIR706.codA. Transformants were seen after selection on kanamycin supplemented media. Eight putative transformants were picked and restreaked twice onto BG-11 supplemented with 200 µg/ml kanamycin. A PCR performed with primers IR706F and IR706R confirmed the integration of *codA* in all eight putative transformants at the targeted neutral site to produce homoplasmic 6803.IR706.codA strains (Figure 5.12). Transformants 6803.IR706.codA_3, 6803.IR706.codA_4 and 6803.IR706.codA_6, were selected for further analysis.

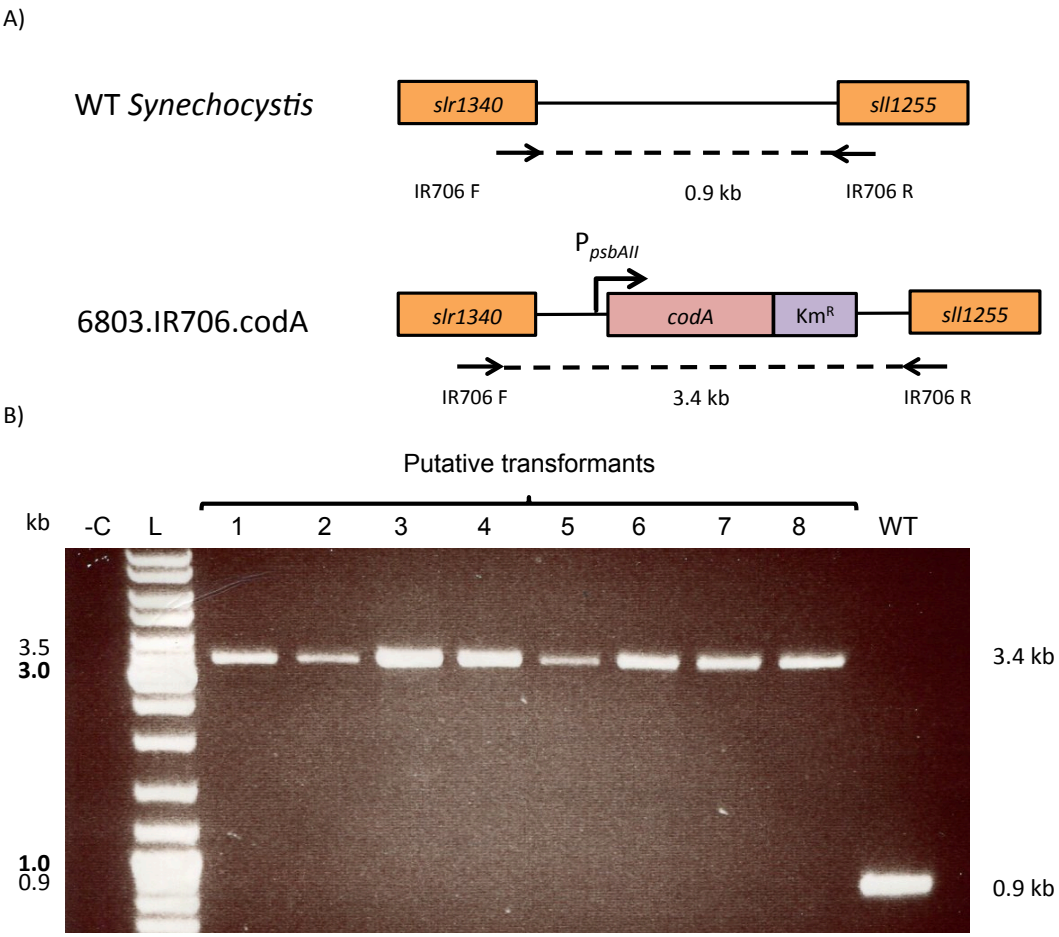


Figure 5.12 PCR screening for successful transformants in *Synechocystis* containing *codA* at the chosen neutral site.

(A) Schematic showing the binding sites and expected fragments sizes for WT *Synechocystis* and 6803.IR706.codA transformants. (B) PCR screening of putative transformants of 6803.IR706.codA. WT *Synechocystis* (WT) and no DNA template (-C) were used as controls. Homoplasmic strains confirmed by the single band at 3.4 kb.

5.2.3.3 Western blot analysis confirms expression of *codA* in 6803.IR706.codA strains

When Western blot analysis with anti-HA tag antibodies was performed on the crude extract of three transformants of 6803.IR706.codA, the expression of *codA* was detected in all transformants (Figure 5.13).

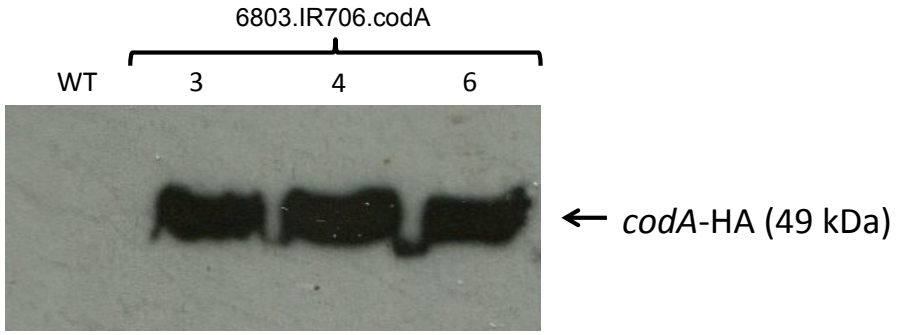


Figure 5.13 Western blot analysis showing the presence of HA tagged codA gene in successful 6803.IR706.codA transformants

Western blot analysis was performed with samples of 6803.IR706.codA using the ECL method with anti-HA antibodies and ECL anti-rabbit IgG antibodies. WT *Synechocystis* (WT) was used as a negative control. The expected band size of CodA is 49 kDa, as seen in all three 6803.IR706.codA transformants.

5.2.3.4 Growth spot test confirms that expression of *codA* confers sensitivity to 5-FC

The Western blot analysis (Figure 5.13) confirmed the expression of *codA* at the neutral site under the control of the P_{psbAII}. In order to confirm that this expression resulted in sensitivity to 5-FC a growth 'spot' test was performed. The three 6803.IR706.codA transformants were able to grow on BG-11 but were unable to grow on BG-11 supplemented with various concentrations of 5-FC (0.25 and 0.5 mg/ml) (Figure 5.14).

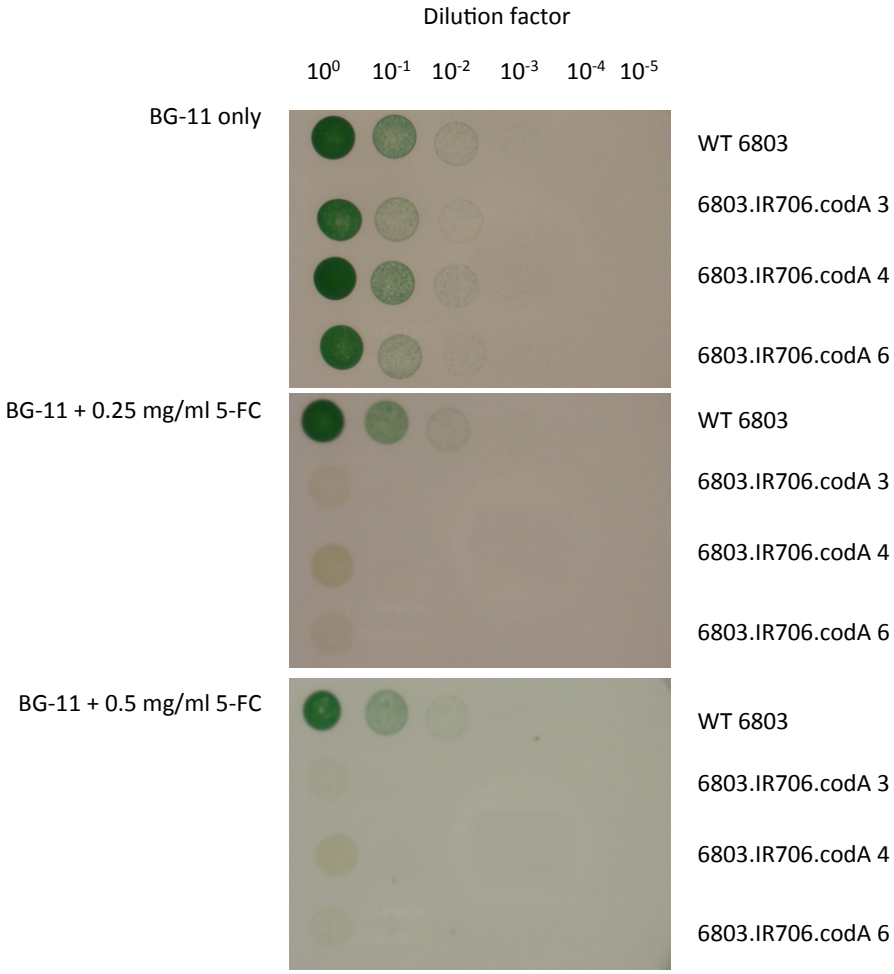


Figure 5.14 Growth 'spot' test confirms expression of *codA* in successful transformants.

A 10-fold serial dilution was performed on a culture of WT *Synechocystis* (WT 6803) and three successful 6803.IR706.codA transformants. These strains were then grown on BG-11 and BG-11 supplemented with 0.25 and 0.5 mg/ml 5-FC. Photographs were taken after 5 days of incubation. Growth of the three 6803.IR706.codA transformants was inhibited by the presence of 5-FC.

5.2.4 Developing a protocol for producing marker-less transformants with the chassis strain, 6803.IR606.codA

The selection protocol used in the generation of marker-less transformants when using negative selectable markers such as *sacB* differs to the protocol when using positive markers such as antibiotic resistance markers (Liu et al., 2011). When selecting for transformants using an antibiotic resistance marker, putative transformants that contain at least one copy of the resistance gene are able to survive. Further selection can be performed on these colonies once selected to ensure all copies of the genome have integrated the resistance gene. However, with negative selection, if a single copy of the negative selectable marker remains, cell death will occur when the selective pressure is applied, there is no opportunity to lose the remaining copies of the marker. Therefore

complete segregation of the mutant alleles needs to be achieved before the selective pressure can be applied.

The protocols used to produce marker-less transformants using the *sacB* – sucrose method in *Synechocystis*, were therefore used as a template for developing a protocol to produce marker-less transformants using *codA* – 5-FC.

5.2.4.1 Unsuccessful attempts at producing marker-less transformants using 6803.IR706.codA

Numerous attempts were made to transform 6803.IR706.codA with pIR706.cat to produce marker-less transformants containing *cat* at the neutral site. One attempt led to the production of heteroplasmic strains of *Synechocystis* that contained copies of both the *cat* gene and *codA* gene (Figure 5.15). The protocol used by Liu *et al*., to produce marker-less transformants in *Synechocystis* using the *sacB* – sucrose method, was used on this occasion (Liu et al., 2011). A single 6803.IR706.codA strain (6803.IR706.codA _6) was grown to an OD_{730} of 1.5. Cells were prepared in the same way as when transforming WT *Synechocystis* (see 2.4.13 or details). The amount of cells used in each transformation mix differed between the two methods, instead of using 200 μl, 10 μl was used for each transformation mix. Approximately 0.4 μg of pIR706.cat was added before the mixture was incubated at 30°C for 4 hours. To each mixture 2 ml of BG-11 was added. The tubes were incubated at 25°C in a shaking incubator for 4 days before 1 ml and 0.5 ml was plated onto BG-11 supplemented with 0.25 mg/ml 5-FC. These plate were then incubated at 30°C.

Colonies appeared after 6 days. Thirty colonies were picked and restreaked onto BG-11 supplemented with 0.25 mg/ml 5-FC. After 2 days, eight putative transformants began to grow. A PCR using primers IR706F and IR706R was performed on the DNA of these putative transformants. The results suggested no marker-less transformants were produced; however, some transformants (# 2, 8, 10, 12) appeared to have copies of both the *codA* gene and the *cat* gene in their genome, i.e. their genomes were heteroplasmic (Figure 5.15).

These four heteroplasmic transformants, along with putative transformant 18 were restreaked onto BG-11 medium containing 0.5 mg/ml 5-FC in an attempt to select for the loss of the remaining copies of *codA*. After restreaking the heteroplasmic transformants twice on 0.5 mg/ml 5-FC, a PCR showed that only one (number 10) of the four 'heteroplasmic' strains still contained *cat* (Figure 5.15). The presence of *cat* in the remaining 'heteroplasmic' strain was then investigated by streaking the strain onto BG-11

supplemented with chloramphenicol (34 μg/ml), however the strain did not grow suggesting the absence of the *cat* gene in this transformant.

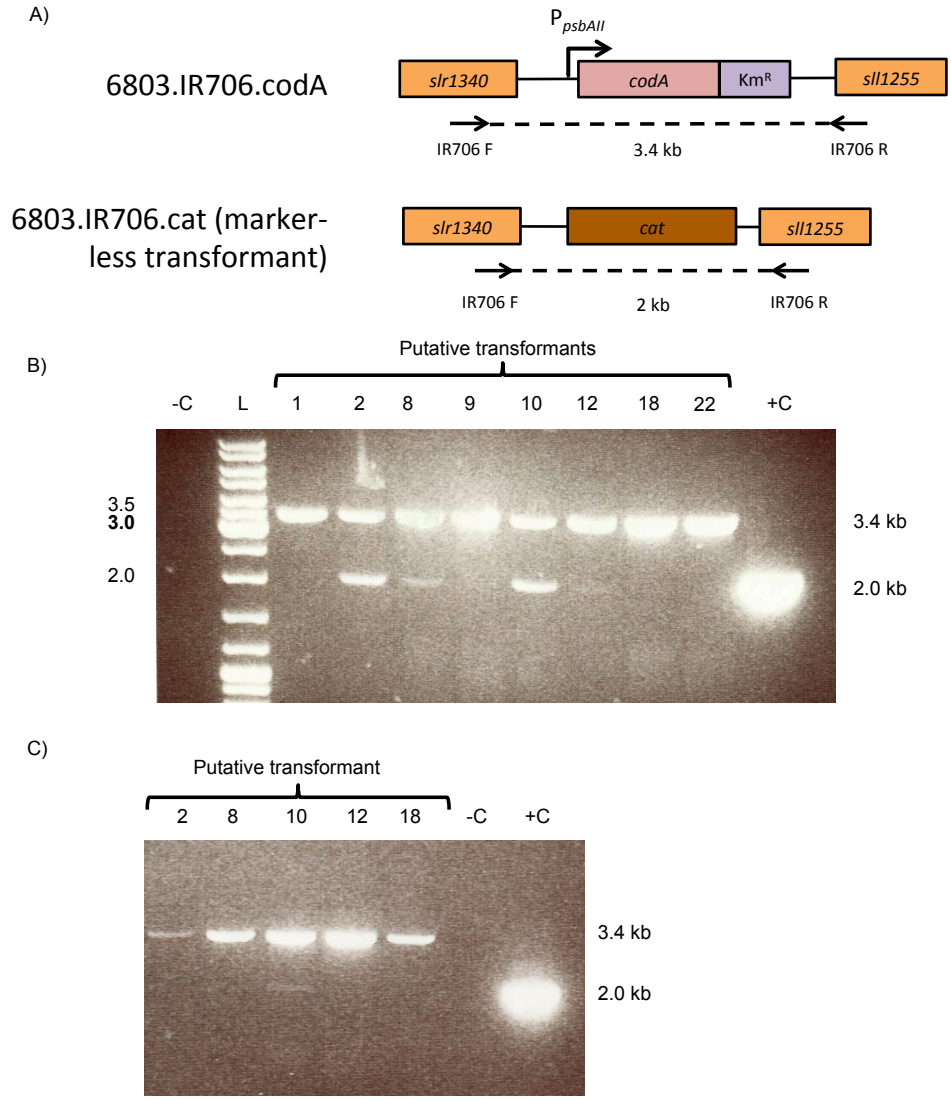


Figure 5.15 PCR screening of putative marker-less transformants containing both copies of *codA* and *cat*.

(A) Schematic of the primer binding sites and the expected sizes of 6803.IR706.codA and 6803.IR706.cat (marker-less transformants). (B) PCR screening of putative transformants. 6803.IR706.codA (+C) and no DNA template (-C) used as controls. Putative transformants 2, 8, 10, 12 appear to have copies of *cat* integrated between *slr1340* and *sll1255*, but copies of *codA* remain, shown by the band at 3.4 kb. (C) Another PCR screen after further selection on 5-FC supplemented media. Same controls were used, along with putative transformant 18. Only putative transformant 10, remained heteroplasmic, the other putative transformants appear to have lost the *cat* gene.

5.2.4.2 An extension of the region of homology results in marker-less transformants

As attempts to produce marker-less transformants with pIR706.cat were unsuccessful, the idea of extending the region of homology in the transforming plasmid, pIR706.cat, was

considered. There is some evidence to suggest that increasing the region of homology increases the transformation efficiency (Zang et al., 2007). Currently the regions of homology between 6803.IR706.codA and pIR706.cat are 445bp and 330bp but by increasing the regions of homology to ~1000bp, ideally the transformation efficiency will increase, along with the likelihood of producing of marker-less transformants.

5.2.4.2.1 Creation of pIR706.Ex and pIR706.Ex.Cat to test the production of marker-less transformants

Primers, IR706.ExF and IR706.ExR, were used to amplify approximately 1000 bp either side of the NheI restriction sites in the neutral site to generate IR706.Ex (Figure 5.16). The 2 kb fragment was then cloned into the pJET1.2/blunt cloning vector. A PCR using the primers IR706.ExF and IR706.ExR and DNA sequencing with pJET F and pJET R primers confirmed the creation of pIR706.Ex (Figure 5.16).

The *cat* gene from pARG1.3 was digested with EcoRV and ZraI and cloned into pIR706.Ex at the HpaI site to generate pIR706.Ex.cat. A PCR performed with primers IR706.ExF and IR706.ExR alongside sequencing with primers pJETF, pJETR, IR706F and IR706R confirmed the creation of pIR706.Ex.cat.

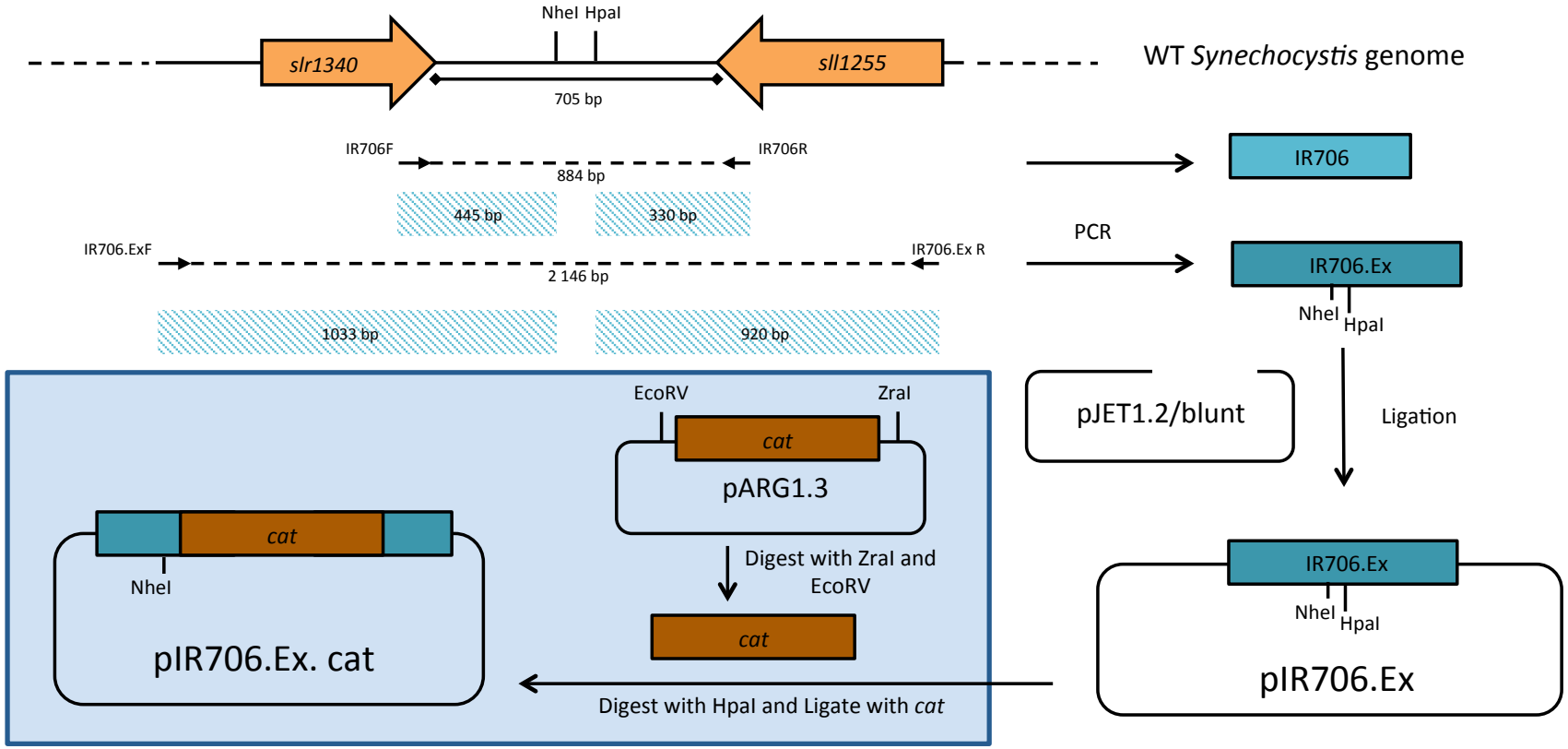


Figure 5.16 Diagram describing the extension of the region of homology and the construction of the plasmids with the extended region of homology.

Primers IR706.ExF and IR706.ExR were designed to amplify approximately 1000bp of homology either side of the gene insertion site. The extended region of homology (IR706.Ex) was then used to construct pIR706.Ex and pIR706.Ex.cat required to produce marker-less transformants.

5.2.4.3 Establishment of a method for the successful production of marker-less strains of *Synechocystis*

A similar protocol to one that produced the 'heteroplasmic strains' (5.2.4.1) was employed. A single 6803.IR706.codA transformant, 6803.IR706.codA_3, was grown to an OD_{730} of 0.85. To 10 μl of cells, 1 μg of pIR706.Ex.cat was added. A control mixture, where no DNA was added to the cells was also prepared. The cells were then incubated at 30°C for 5 hours before 2 ml of BG-11 medium was added to each transformation mix. The tubes were then placed in a 30°C shaking incubator for 4 days before 1ml and 0.5 ml of the transformation mix was plated onto BG-11 supplemented with 0.5 mg/ml 5-FC. The plates were then incubated at 30°C.

After 7 days colonies appeared, 16 putative transformants were picked and restreaked onto BG-11 supplemented with 0.5 mg/ml 5-FC. From the 16 putative transformants 11 grew after 2 days. PCR using primers IR706.Ex.F and IR706.Ex.R was performed on 5 of these putative transformants; one transformant (#13) appeared to have lost the *codA* gene and gained the *cat* gene (Figure 5.17). The other putative transformants, which later grew on 5-FC, were tested by PCR and another two transformants (#5 and #15) appeared to have replaced *codA* with *cat* (Figure 5.17). One line (#3) appeared to have reverted back to WT *Synechocystis* (3) and another (#14) produced a band at ~2.8kb, an unexpected size (Figure 5.17). Other putative transformants still contained the *codA* gene.

It should be noted that there were a significant number of colonies present on the control plates, where no DNA was added to 6803.IR706.codA. The number of colonies on the control plates was approximately half the number of colonies on the plates with the putative transformants.

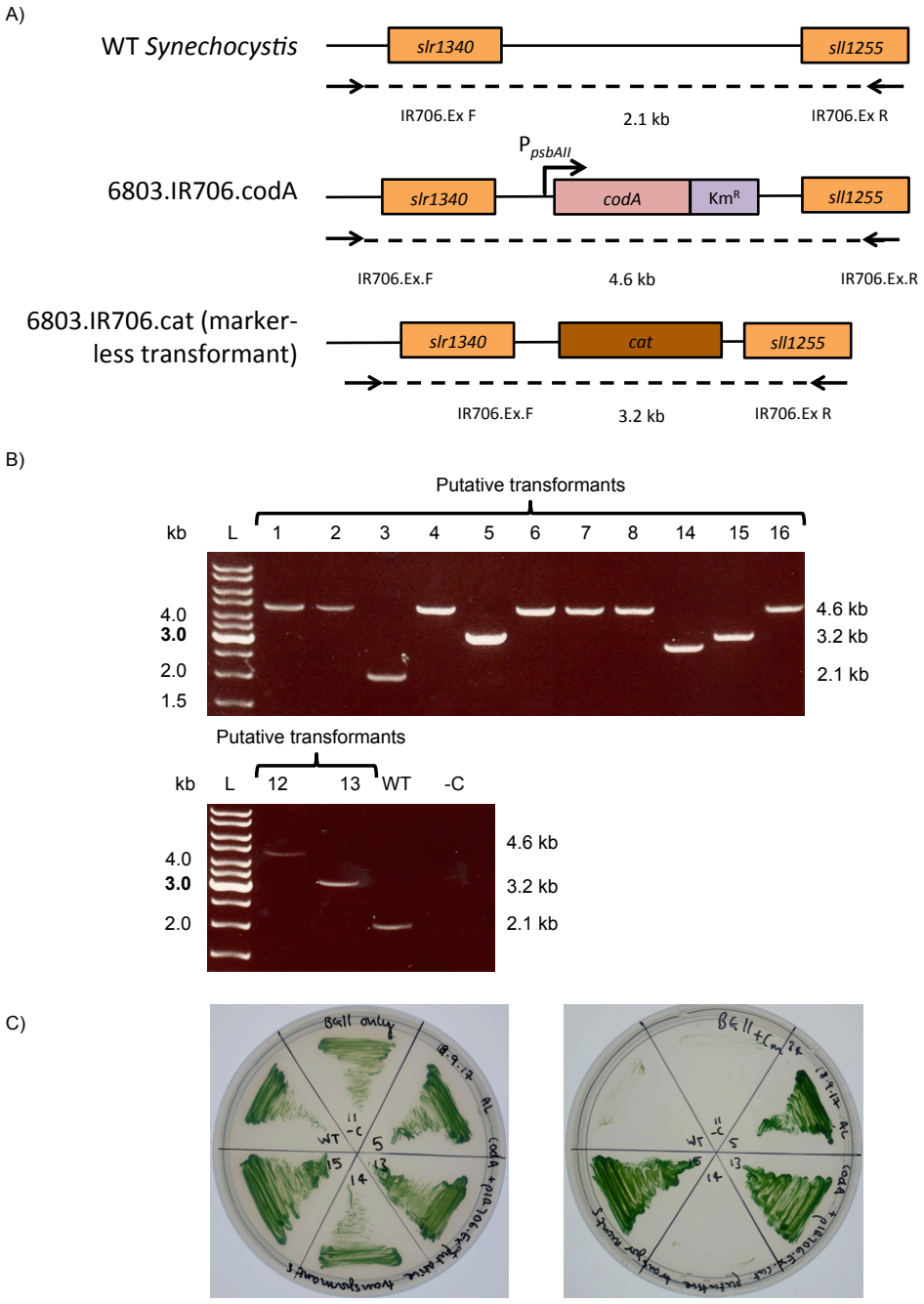


Figure 5.17 Screening of successful marker-less transformants that replaced *codA* with cat.

(A) Schematic of the primer binding sites and the expected sizes of WT *Synechocystis*, 6803.IR706.codA and 6803.IR706.cat (marker-less transformants). (B) PCR screening of putative transformants. WT *Synechocystis* (WT) and no DNA template (-C) were used as controls. Putative transformants 5, 15 and 13 appeared to be homoplasmic transformants containing *cat* with a single band at 3.2 kb. (C) Photographs of WT *Synechocystis* and the putative transformants (same as the ones used in the PCR) 5, 11 (-C), 13, 14 and 15 streaked on BG-11 and BG-11 supplemented with 34 µg/ml chloramphenicol to confirm expression of *cat*. Photograph taken after 4 days of incubation.

To confirm the replacement of *codA* with *cat,* lines #5, #13, and #15 were picked and restreaked onto BG-11 and BG-11 that was supplemented with 34 μg/ml chloramphenicol. Line #11, which still contained the *codA* gene (not seen in Figure 5.17) and WT *Synechocystis* were grown as controls. Line #14 was also grown due to the similar size product to marker-less transformants. Only lines #5, #13 and #15 could grow in the presence of chloramphenicol, confirming the integration of *cat* in these transformants to produce 6803.IR706.cat. The production of 6803.IR706.cat was also confirmed through DNA sequencing analysis of the *cat* PCR product using primers IR706F and IR706R.

The production of 6803.IR706.cat confirms that the production of marker-less transformants using 5-FC to select for the loss of *codA* is possible. To confirm that the production of marker-less transformants was reproducible, another attempt using the method described above showed that adding 2 μg of DNA and shaking the cells at 25°C instead of 30°C for four days also resulted in the production of at least four successful transformants of 6803.IR706.cat (results not shown). As before there was a high number of colonies on the control plate: the number of colonies on control plate (30) when 0.5 ml was plated was half the number of putative transformants (67). When 1 ml is plated the number of colonies on the control plate (108) is similar to the number of putative transformants (115). Of the six putative transformants tested only two still contained *codA*. In the previous test, of the 16 tested, 13 contained the *codA* gene. The ability to grow on 5-FC is most likely due to a mutation in *codA* that overcomes sensitivity towards 5-FC.

Using pIR706.Ex.cat rather than pIR706.cat, increasing the region of homology, may have had the desired effect of producing of marker-less transformants, however as other changes were made to the protocol, it may not be the only factor in the success of the method.

The production of marker-less transformants using pIR706.cat and pIR706.Ex.cat was made slightly more difficult due to the construction of the transforming plasmid. The *cat* gene was cloned at the HpaI site of the transforming plasmid; however, *codA* was located at the NheI site. This misalignment results in three regions of homology between IR706 (Figure 5.18). If homologous recombination does occur, marker-less transformants, with *cat* replacing *codA,* may not be the only product. A strain of *Synechocystis* containing both *codA* and *cat* at the neutral site could be produced; however, selection with 5-FC should ensure this transformant is selected against (Figure 5.18). A strain that no longer contains *codA* or *cat* that appears as WT *Synechocystis* could also be produced and would be present after selection with 5-FC (Figure 5.18).

From the PCR results other products, not mentioned above, could also be produced, e.g. line #14 produced a band at ~2.8 kb, which does not correlate to any of the possible homologous recombination events mentioned above (Figure 5.18).

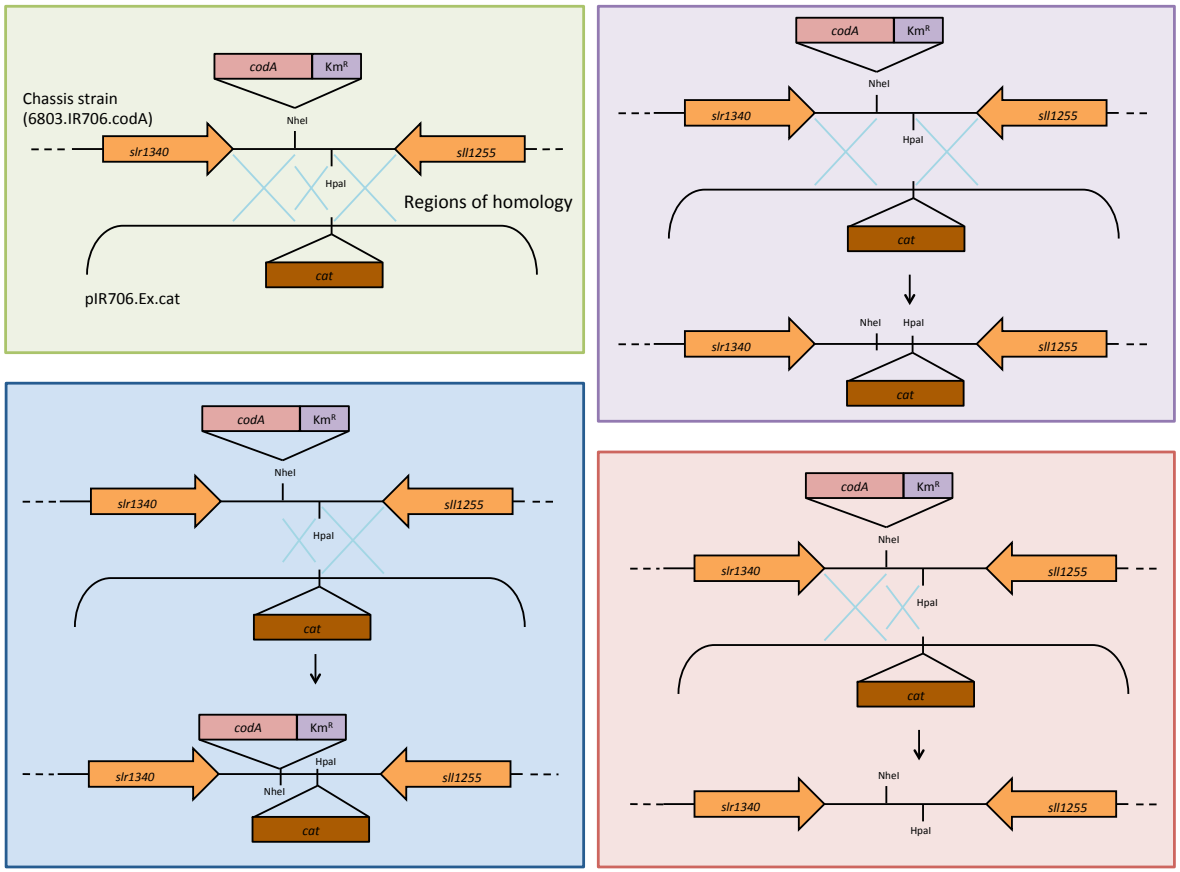


Figure 5.18 Diagram describing the various homologous recombination events that are possible due to the construction of the pIR706.Ex.cat plasmid.

There are three regions of homology that can lead to different homologous recombination events. For the production of marker-less transformants, the insertion of *cat* (purple box) is required. Transformants with both *codA* and *cat* are possible (blue box), as are transformants with neither *codA* nor *cat* (red box).

5.2.5 Introduction of the limonene synthase gene at the neutral site using the marker-less transformation method

Demonstrating that the *codA* – 5-FC method can be used to introduce another gene, one that is not an antibiotic resistance gene, would add further evidence that the *codA* - 5-FC method can produce marker-less transformants. The synthetic limonene synthase gene, *LS*, described in Chapter 3 was introduced at the neutral site using the same protocol used to introduce *cat* into *Synechocystis*.

5.2.5.1 Creation of pIR706.Ex.LS

Primers, AII.LS_F and AII.LS_R were used to amplify *LS* under the control of the P_{psbAII} from pLAH.AII.LS. The primers introduced an NheI restriction site upstream of the promoter and a HpaI restriction site downstream of *LS*. By cloning the gene between the two sites, this potentially avoids the unwanted homologous recombination events that occurred with pIR706.Ex.cat (Figure 5.19). PCR using AII.LS_F and AII.LS_R successfully amplified *LS* under the control of the P_{psbAII}.

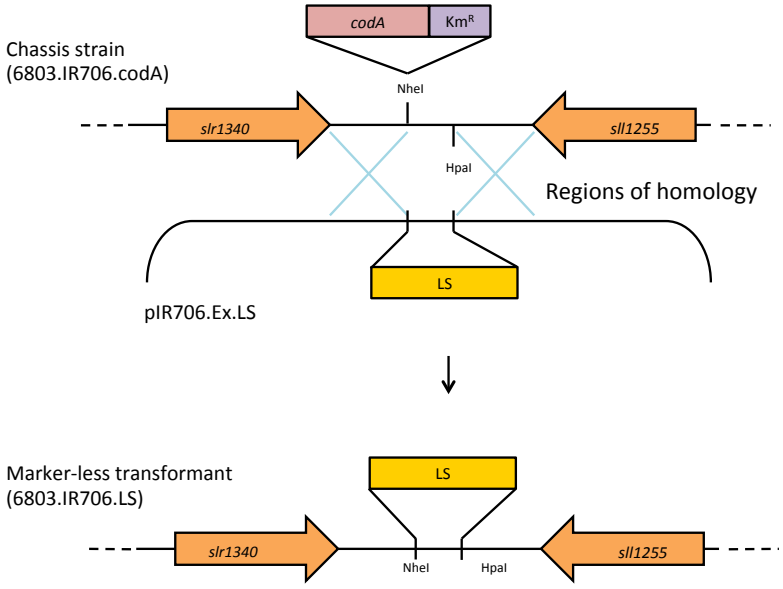


Figure 5.19 Diagram demonstrates insertion of *LS* between the HpaI and NheI restriction sites.

With only two regions of homology, the replacement of *codA* and *Km^R* with *LS* should be the only possible gene arrangement in transformants.

The PCR fragment was digested with NheI and HpaI and cloned into the restriction sites in pIR706.Ex. A PCR using IR706F and IR706R and DNA sequencing with primers IR706F, IR706R and LS.midF confirmed the construction of pIR706.Ex.LS (Figure 5.20).

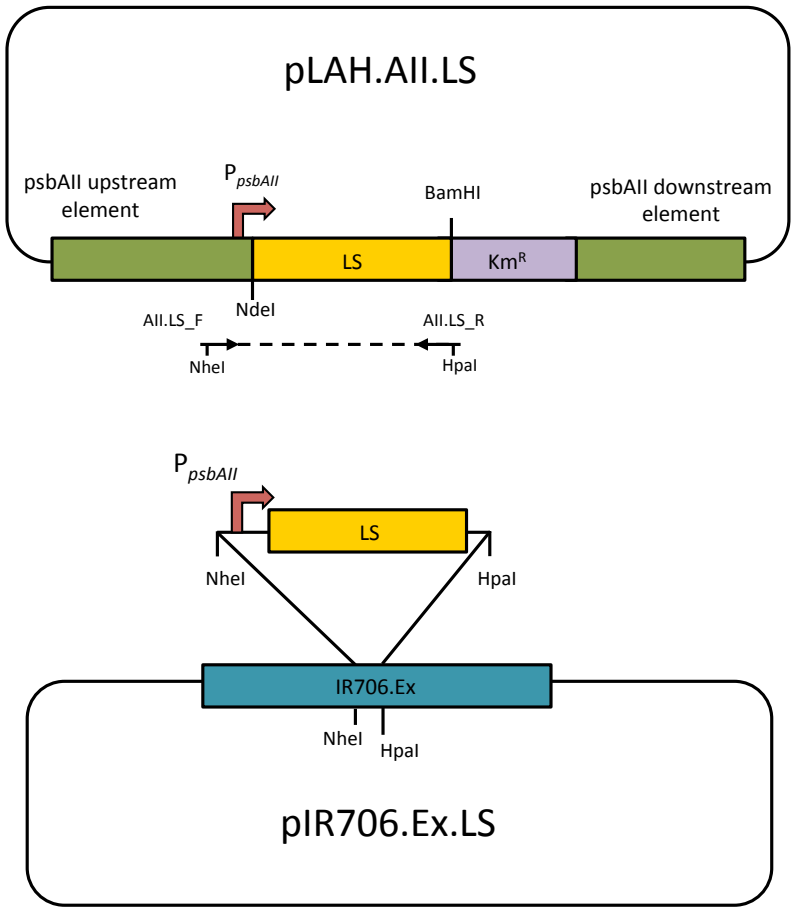


Figure 5.20 Method used to produce pIR706.Ex.LS

The *LS* gene under the control of the P_{psbAII} promoter was amplified from pLAH.AII.LS and cloned into pIR706.Ex.LS.

5.2.5.2 Transformation of pIR706.Ex.LS with 6803.IR706.codA produces marker-less transformants containing the limonene synthase gene

Three 6803.IR706.codA (#3,4 and 6) transformants were grown to an OD_{730} between 0.5-0.8 to check they could all produce marker-less transformants. The *codA*-5-FC method established (see 2.4.14 for details) was used to transform 6803.IR706.codA with pIR706.Ex.LS, although after the addition of DNA cells were incubated for 8 hours at 30°C.

Colonies appeared after 9 days, and 16 colonies were picked from each 6803.IR706.codA transformation (48 colonies in total) and restreaked onto BG-11 supplemented with 0.5 mg/ml 5-FC. A PCR with IR706.Ex.Flank.F and IR706.Ex.Flank.R confirmed that of the 48 colonies picked, 14 had replaced *codA* with *LS* to produce 6803.IR706.LS (Figure 5.21). Marker-less transformants were produced from each 6803.IR706.codA transformant tested (Figure 5.21). Successful transformants from each 6803.IR706.codA transformation mix were chosen for further analysis. Table 5.2 summarises the results of the PCR on all 48 putative transformants. Almost three-quarters of the putative transformants still contain

the *codA* gene. A few transformants produced products that did not correlate with the expected result (Table 5.2).

Controls for each 6803.IR706.codA transformation, where no DNA is added, were prepared. Table 5.3 summarises the number of colonies that were present on these plates after 9 days. Similar to the experiments involving the production of marker-less transformants containing *cat*, there are a significant number of colonies present on the control plates.

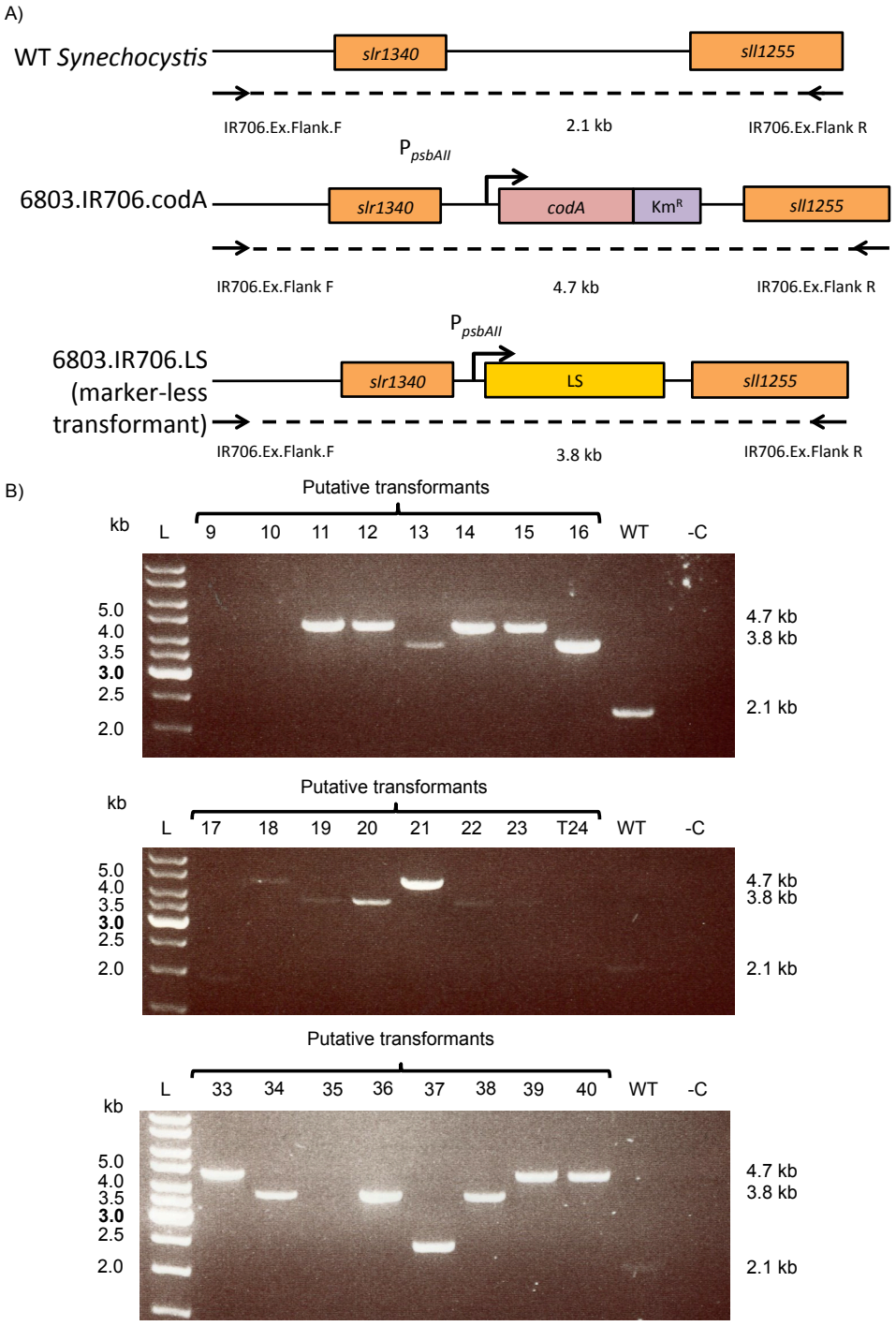


Figure 5.21 PCR Screening for marker-less transformants of *Synechocystis* containing *LS*.

(A) Schematic showing the primer binding site and expected fragments of WT *Synechocystis*, 6803.IR706.codA and 6803.IR706.LS. (B) PCR screening of successful transformants containing *LS*. A single band at 3.8 kb indicates homoplasmic strains containing *LS*. Putative transformants 1-16 are from the transformation of 6803.IR706.codA_3, 17-32 from 6803.IR706.codA_4 and 33-48 from 6803.IR706.codA_6. WT *Synechocystis* (WT) and no DNA (-C) were used as controls.

Table 5.2 Summary of all the PCR screening results for marker-less transformants containing *LS*.

After screening a total of 48 putative transformants (16 from each of the 6803.IR706.codA transformation mixes), 14 appear to be marker-less transformants.

Band size (kb)	Putative transformants number	Total number of transformants	Expected transformant strain
1.9	17	1	Unknown
2.4	37, 41	2	Unknown
3.8	13, 16, 19, 20, 22, 23, 26, 27, 29, 31, 32, 34, 36, 38	14	6803.IR706.AII.LS
4.7	1-12, 14, 15, 18, 21, 24, 25, 28, 30, 35, 39, 40, 42- 48	31	6803.IR706.codA

Table 5.3 Summary of the number of colonies present after transformation of 6803.IR706.codA with pIR706.Ex.LS

The number of colonies present on the control plates for the majority of the transformation reactions is greater than half the number of colonies present on the plates containing the putative transformants.

Transformation mixture	Volume plated (ml)	Number of colonies
6803.IR706.codA_3 + no DNA	0.5	35
	1	75
6803.IR706.codA_3 + pIR706.Ex.LS	0.5	52
	1	90
6803.IR706.codA_4 + no DNA	0.5	22
	1	54
6803.IR706.codA_4 + pIR706.Ex.LS	0.5	31
	1	62
6803.IR706.codA_6 + no DNA	0.5	40
	1	54
6803.IR706.codA_6 + pIR706.Ex.LS	0.5	61
	1	113

5.2.5.3 Confirming the absence of kanamycin resistance gene

The marker-less selection method is based on the loss of both *codA* and *nptI*, and their replacement with the G.O.I. Consequently, successful transformants should be kanamycin sensitive as well at 5-FC resistant. To test this, three selected 6803.IR706.LS transformants (#13, 19 and 34) and the 6803.IR706.codA strains were grown on BG-11, BG-11 supplemented with kanamycin and BG-11 supplemented with 0.5 mg/ml 5-FC. As expected the growth 'spot' tests demonstrated that 6803.IR706.codA strains were able to grow in the presence of kanamycin but unable to grow in the presence of 5-FC (Figure 5.22). The growth of the 6803.IR706.LS transformants in the presence of 5-FC and inability to grow on kanamycin confirm the loss of the *codA*/Km^R cassette (Figure 5.22). These results along with the PCR results confirm the production of marker-less transformants with the limonene synthase gene at the neutral site.

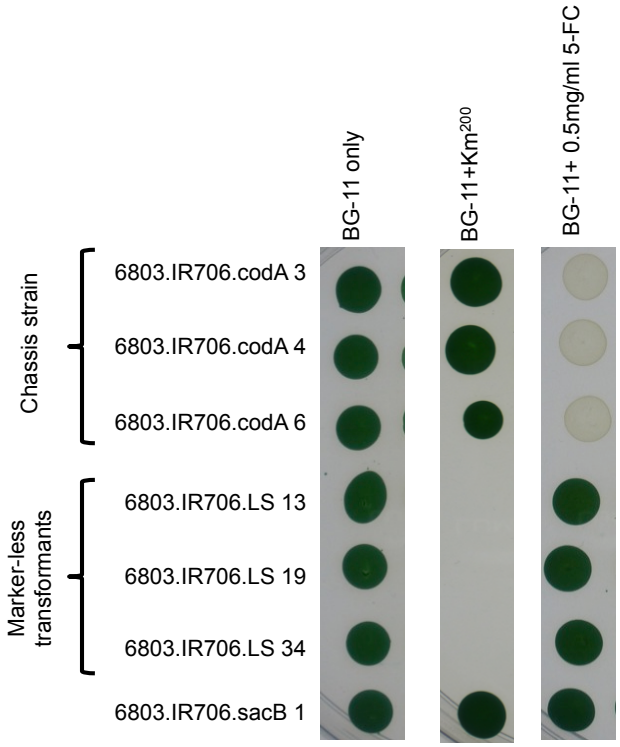


Figure 5.22 Growth 'spot' test confirms the loss of *codA* and kanamycin resistance gene in successful 6803.AII.LS transformants.

Three 6803.IR706.codA and three 6803.IR706.LS transformants were grown on BG-11, BG-11 supplemented with 200 µg/ml kanamycin and BG-11 supplemented with 0.5 mg/ml 5-FC. 6803.IR706.sacB_1 was used as a control. 6803.IR706.codA was able to grow on kanamycin but unable to grow on 5-FC, whilst the 6803.IR706.LS were unable to grow on kanamycin but were able to grow on 5-FC.

5.2.5.4 Expression of limonene synthase is detected by Western blot analysis

To confirm the expression of LS, Western blot analysis was performed on the crude extract of three 6803.IR706.LS (#13, 19 and 34) transformants. LS was detected using anti-HA antibodies in all three 6803.IR706.LS transformants with a band at 65 kDa present that is absent in the negative control, 6803.AII, and present in the positive control, 6803.AII.LS (Figure 5.23).

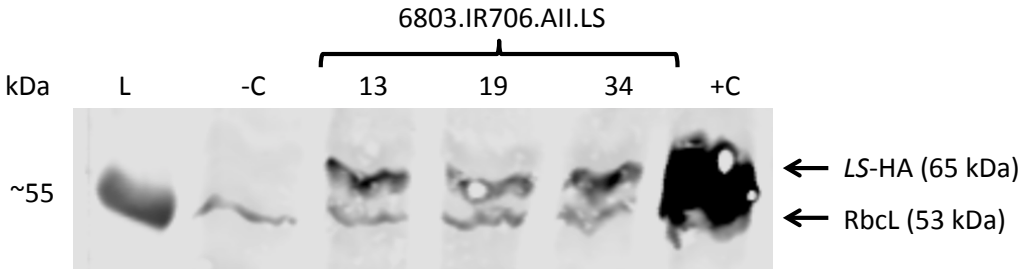


Figure 5.23 Western blot analysis confirms the expression of LS in marker-less transformants of *Synechocystis*.

The Western blot was performed with whole cell extracts and anti-HA and anti-RbcL antibodies, using the Odyssey® Infrared Imaging System for detection. The samples were run on a 10% SDS-PAGE gel. The expected band size for LS is 65 kDa and 53 kDa for RbcL, which acted as loading control. Protein size was determined using the PageRuler™ Prestained Protein ladder (Thermo Scientific), lane L. Strains 6803.AII (-C) and 6803.AII.LS (+C), were grown as controls.

5.2.6 Return to *sacB*: expression of *sacB* confers sensitivity to sucrose.

After receiving information on how the sucrose medium should be prepared, the *sacB* – sucrose method was revisited to investigate why both 6803.IR706.sacB and 6803.IR706.sacB' strains did not confer sensitivity to sucrose.

5.2.6.1 Growth 'spot' test analysis of *sacB*

The lack of mutations in the *sacB* gene, suggested that there must be another explanation as to why *sacB* expression did not confer sensitivity. After obtaining a clear protocol for using the *sacB* – sucrose method (kindly provided from David Lea Smith, University of Cambridge), differences with the preparation of the media used in the earlier growth 'spot' test (5.2.1.6) were observed. In the protocol used by the Cambridge Lab, $NaHCO_3$ is not present in the BG-11 medium with sucrose and the 5% (w/v) sucrose is filter sterilised, not autoclaved.

Growth 'spot' tests were repeated using a new medium based on the protocol provided. Three 6803.IR706.sacB transformants, WT *Synechocystis* and 6803.AII (used as controls) were grown on BG-11 (also prepared without $NaHCO_3$), BG-11 with kanamycin and BG11 with 5% and 7.5% (w/v) sucrose. The three 6803.IR706.sacB transformants were unable to grow on 5% or 7.5% (w/v) sucrose (Figure 5.24) but were able to grow on BG-11 and BG-11 with kanamycin. WT *Synechocystis* and 6803.AII grew on the plate containing BG-11 with 5% (w/v) sucrose but struggled to grow on medium containing 7.5% (w/v) sucrose.

Another growth 'spot' test was performed with three 6803.IR706.sacB' transformants, WT *Synechocystis* and 6803.AII (used as controls) on BG-11, BG-11 supplemented with 5%

(w/v) sucrose and BG-11 supplemented with kanamycin. The results of these 'spot' tests demonstrated that the expression of *sacB* in the presence of sucrose does lead to the cell death in *Synechocystis* as expected. The *sacB* - sucrose method can therefore be used with 6803.IR706.sacB and 6803.IR706.sacB' to produce marker-less transformants.

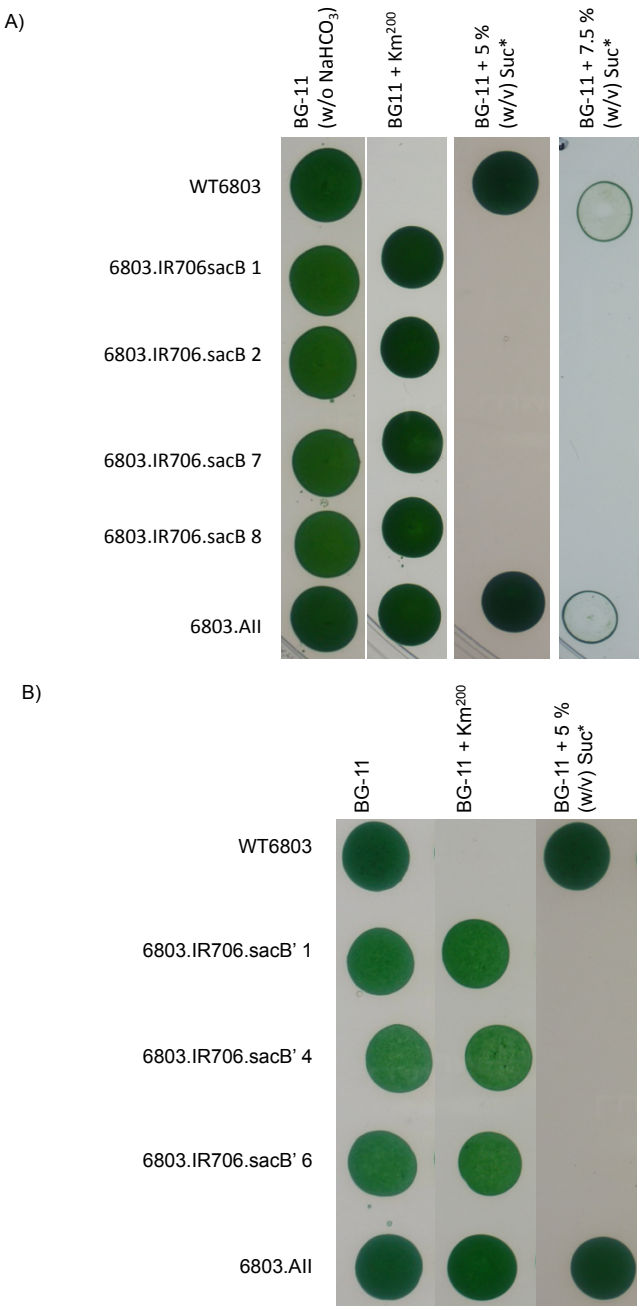


Figure 5.24 Growth spot tests testing the expression of *sacB* in 6803.AII.sacB (A) and 6803.IR706.sacB' in the presence of sucrose.

(A) Four transformants of 6803.IR706.sacB were grown on BG-11 (w/o $NaHCO_3$), BG-11 supplemented with 200 μg/ml kanamycin and BG-11 (w/o $NaHCO_3$) supplemented with 5% and 7.5% (w/v) sucrose. 6803.IR706.sacB strains were able to grown on kanamycin but were unable grow on media supplemented with 5% sucrose. All strains were unable to grow on media supplemented with 7.5 % sucrose. In both (A) and (B) WT *Synechocystis* (WT 6803) and 6803.AII were grown as controls. Photographs were taken after 7 days incubation. (B) Three transformants of 6803.IR706.sacB' were grown on BG-11, BG-11 supplemented with 200 μg/ml kanamycin and BG-11 (w/o $NaHCO_3$) supplemented with 5% (w/v) sucrose. 6803.IR706.sacB' strains were able to grown on kanamycin but were unable grow on media (w/o $NaHCO_3$) and supplemented with 5% sucrose. Photographs were taken after 5 days of incubation.

5.2.6.2 Production of marker-less transformants using the *sacB* selection

All three strains of 6803.IR706.sacB were grown to an OD_{730} between 0.7 and 0.9 and transformed with pIR706.Ex.cat using a similar method used to successfully obtain marker-less transformants using *codA* – 5-FC method (see 2.4.15 for details). After 14 days only two colonies grew. PCR performed on these two putative transformants showed that only one had replaced *sacB* with *cat* to produce 6803.IR706.cat from 6803.IR706.sacB_1 (Figure 5.25). As the other putative transformant was able to grow on sucrose, but still contained *sacB*, this could be due to a mutation in *sacB*.

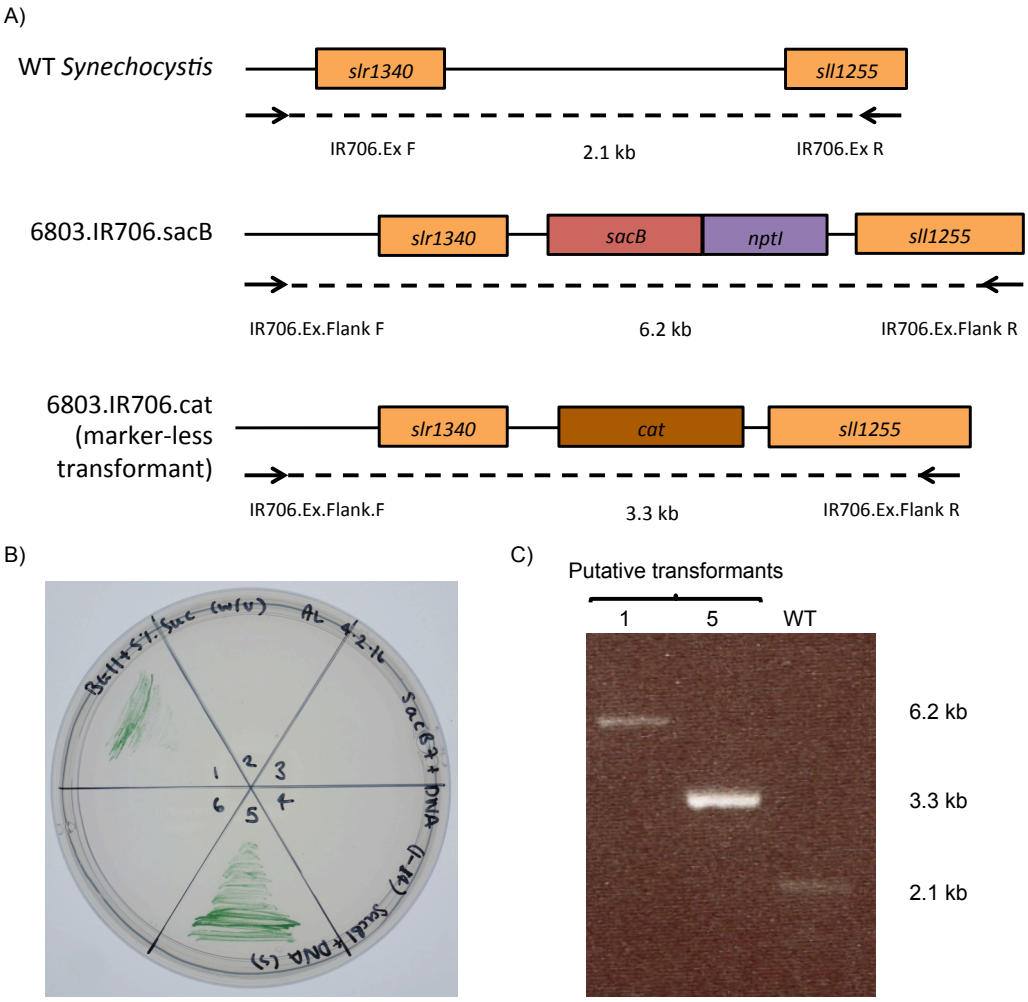


Figure 5.25 Screening of marker-less transformants produced in *Synechocystis* using the *sacB* – sucrose method.

(A) Schematic showing the binding sites and expected fragments sizes of WT *Synechocystis*, 6803.IR706.sacB and 6803.IR706.cat. (B) Photograph of putative transformants picked and restreaked onto BG-11 supplemented with 5% (w/v) sucrose. Only putative transformants 1 and 5 were able to grow. (C) PCR screening of putative transformants confirms the production of marker-less transformant in transformant 5. The strain appears homoplasmic as there is a single band at 3.3 kb.

5.2.7 Comparison of the two marker-less transformation methods

To compare the two methods of producing marker-less mutants, liquid cultures of 6803.IR706.codA_3 and 6803.IR706.sacB_7 were grown to an OD_{730} of approximately 1.2. The *codA*-5-FC method was used for 6803.IR706.codA_3 and *sacB* – sucrose method was used for 6803.IR706.sacB_7. A control transformation, where no DNA was added to either strain, was also prepared.

After 11 days no colonies were seen on the sucrose plates, suggesting transformation into 6803.IR706.sacB_7 was unsuccessful. Colonies were present on all 5-FC plates, including

the no-DNA negative control plates (Table 5.4). The number of colonies on the plate with putative transformants of 6803.IR706.cat is much greater than the number of colonies on the control plate. However, the number of colonies on the control plate is greater than the number of colonies on the plate with putative transformants of 6803.IR706.LS (Table 5.4). Despite this, putative transformants for 6803.IR706.cat and 6803.IR706.LS were picked for screening.

Table 5.4 Summary of the number of colonies from the transformation of 6803.IR706.codA

Colonies were counted after 11 days incubation. * Estimated value.

Transformation mixture	Volume plated (ml)	Number of colonies
6803.IR706.codA + no DNA	0.5	37
	1	33
6803.IR706.codA + pIR706.Ex.cat	0.5	332*
	1	101
6803.IR706.codA + pIR706.Ex.LS	0.5	25
	1	4

To improve the screening process, colonies were picked and suspended in BG-11, before 'spotting' onto BG-11, BG-11 supplemented with 0.5 mg/ml 5-FC and BG-11 with 200 μg/ml kanamycin. Twenty-four colonies of 6803.IR706.codA_3 transformed with pIR706.Ex.LS and 32 colonies of 6803.IR706.codA_3 transformed with pIR706.Ex.cat were screened in this method. After three days, 16 out of 24 putative transformants appear to have replaced *codA* with *LS* (Figure 5.26). A PCR with IR706.Flank.F and IR706.Flank.R on the first six successful transformants confirmed the results from the spot test (Figure 5.26). After seven days all 32 colonies of 6803.IR706.codA_3 transformed with pIR706.Ex.cat grew and showed that 28 out of 32 colonies appear to have lost *codA* and gained the *cat* gene (Figure 5.27). A PCR with IR706.Flank.F and IR706.Flank.R on the first six successful transformants confirmed the results from the spot test (Figure 5.27).

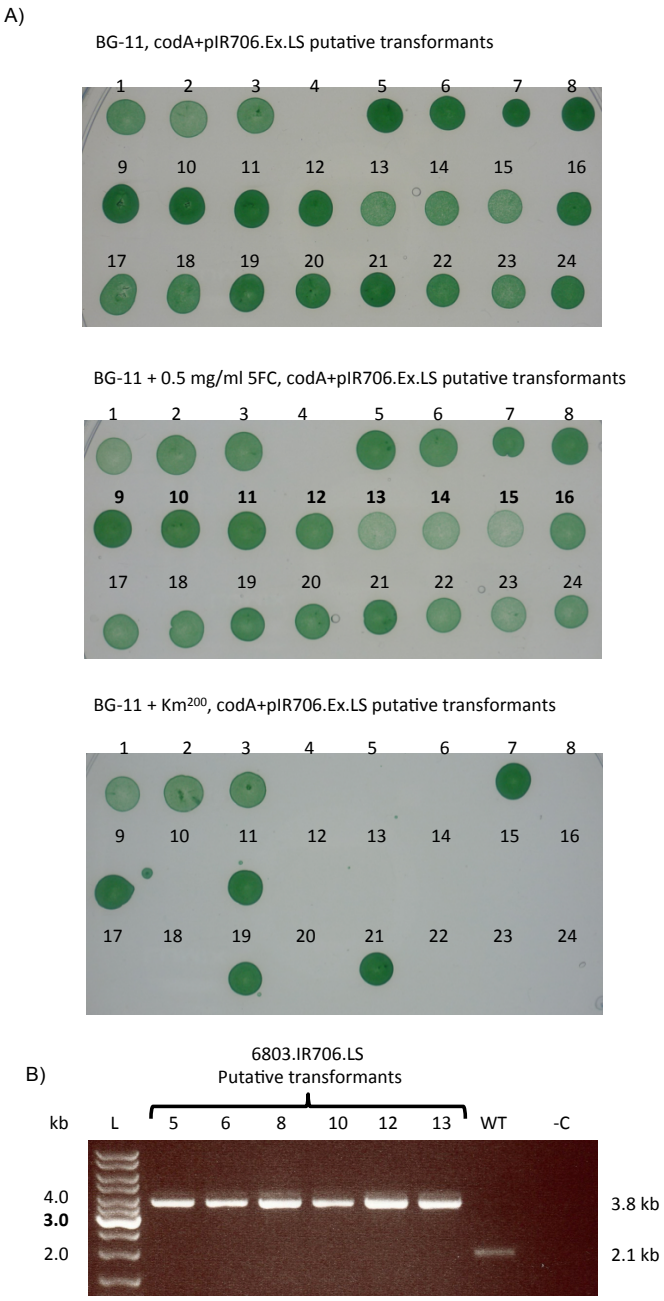


Figure 5.26 Screening of marker-less transformants of *Synechocystis* containing *LS* using the *codA* – 5-FC method.

(A) Growth 'spot' test of 24 putative marker-less transformants grown on BG-11, BG-11 supplemented with 200 µg/ml kanamycin and BG-11 supplemented with 0.5 mg/ml 5-FC. Photographs were taken after 4 days incubation. Of the putative transformants that grew, 15 out of 23 appear to be marker-less transformants. (B) PCR screening with primers IR706 Ex Flank F and IR706 Ex Flank R of the first six putative transformants that appeared to be marker-less transformants from the results of the growth 'spot' tests. Homoplasmic marker-less transformants of *Synechocystis* containing the LS gene are seen with a single band at 3.8 kb, and no band at 4.7 kb. WT *Synechocystis* (WT) and a no DNA template (-C) are used as controls.

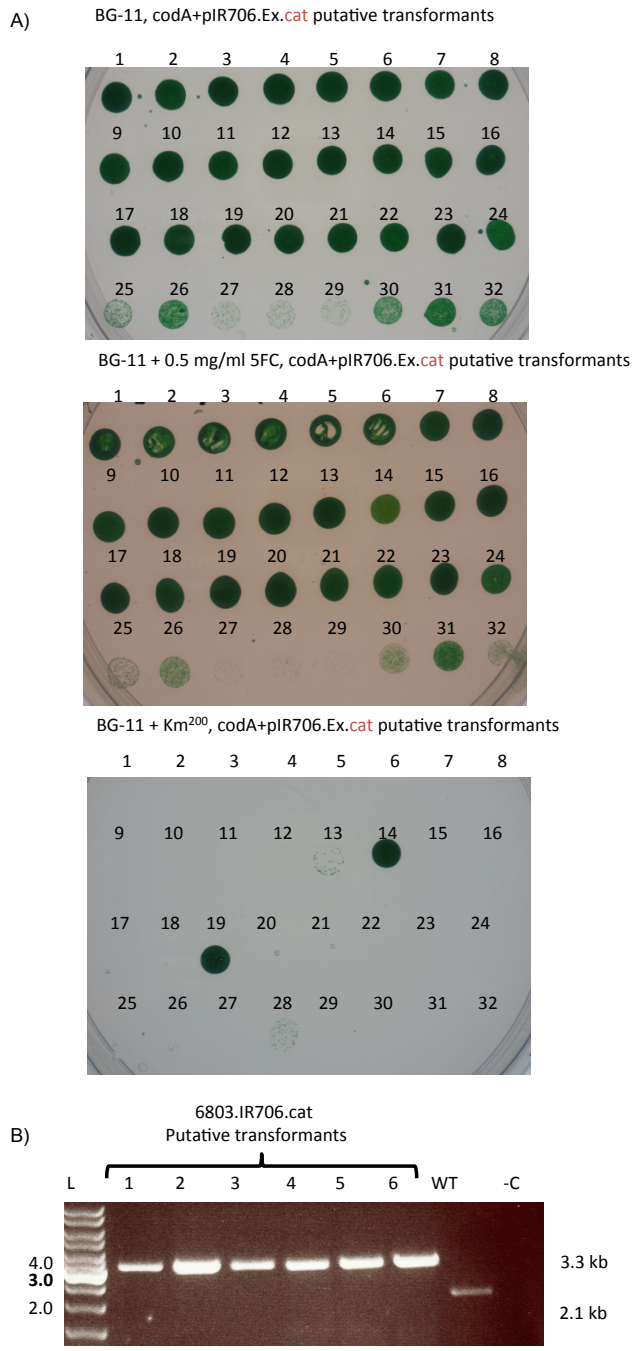


Figure 5.27 Screening of marker-less transformants of *Synechocystis* containing cat using the *codA* – 5-FC method.

(A) Growth 'spot' test of 32 putative marker-less transformants grown on BG-11, BG-11 supplemented with 200 µg/ml kanamycin and BG-11 supplemented with 0.5 mg/ml 5-FC. Photograph taken after 6 days incubation. From the test 28 out of 32 transformants appear to be marker-less transformants containing the *cat* gene. (B) PCR screening with primers IR706 Ex Flank F and IR706 Ex Flank R of the first six putative transformants that appeared to be marker-less transformants from the results of the growth 'spot' tests. Homoplasmic marker-less transformants of *Synechocystis* containing the *cat* gene are seen with a single band at 3.3 kb. WT *Synechocystis* (WT) and a no DNA template (-C) are used as controls.

As the comparison between the two methods was not repeated and no transformants using the *sacB* – sucrose method were produced, a valid comparison between the two methods would be difficult to make. The *sacB* – sucrose method is an established method that others have used to produce marker-less transformants (Lagarde et al., 2000; Lea-Smith et al., 2013; Liu et al., 2011; Xu et al., 2004); however, in all my attempts to use the *sacB* – sucrose method, only a single marker-less transformant was recovered, suggesting that the method is not routine.

5.3 Discussion and Future work

The work presented in this chapter describes the development of a new method for the production of marker-less transformants in *Synechocystis*. The results demonstrate that the expression of a synthetic cytosine deaminase, optimised to accept 5-FC can be used as a method of negative selection in *Synechocystis* when strains are grown in the presence of 5-FC. Furthermore, the expression of the modified *codA* can be used in a similar manner to the *sacB*-sucrose method for the production of marker-less transformants.

The *codA*-5-FC method provides an alternative method to producing marker-less transformants in *Synechocystis* that could potentially be used in other strains that are able to incorporate DNA by homologous recombination such as *Synechococcus* 7942, *Synechococcus* 7002, *Synechococcus* UTEX 2973 and *Anabaena* 7120. The absence of selectable markers in genetically engineered strains removes the potential risk of the antibiotic resistance genes escaping into the environment, and increasing the rate of antibiotic resistance in human diseases. The use of the system also enables multiple rounds of transformations to a strain since the technology is not constrained by the number of selectable markers available.

5.3.1 The *sacB*-sucrose method

From the literature the *sacB*-sucrose method is the most commonly used technique to produce marker-less transformants in *Synechocystis* (Branco dos Santos et al., 2014).

When attempting to use the *sacB*-sucrose method, initially all but one *Synechocystis* strain containing *sacB* appeared to be insensitive to sucrose (Figure 5.5). This result was unexpected, as the same *sacBR*/*nptI* cassette has been demonstrated to be functional in *Synechocystis* (Lea-Smith et al., 2013). After preparing the media differently, all *Synechocystis* strains containing *sacB* were sensitive to sucrose as expected. The unexpected results might be due to the media preparation, as sucrose was autoclaved in the initial growth 'spot' test but sterilised in the successful growth 'spot' tests. At high temperatures, above 110°C, such as those seen during autoclaving, the hydrolysis of sucrose to glucose and fructose occurs, thereby decreasing the concentration of sucrose. The concentration of sucrose may have been too low to have a toxic effect on *Synechocystis*. In the literature, the *sacB*-sucrose method works in *Synechocystis* in the presence of 4.5% (w/v) sucrose, but by autoclaving sucrose the concentration may have fallen below this (Liu and Curtiss, 2009; Liu et al., 2011).

Sodium bicarbonate, $NaHCO_3$, was also absent in the media containing sucrose, when the transformants, 6803.IR706.sacB and 6803.IR706.sacB' were unable to grow. Sodium bicarbonate is usually added to BG-11, to provide an additional carbon source to atmospheric carbon dioxide, although how this might impact on SacB activity is unclear.

Once sensitivity to sucrose was established for the chassis strains of *Synechocystis* these could be used to produce marker-less transformants. However, attempts to the use the *sacB*-sucrose method were not very successful with only a single transformant containing *cat* being produced.

5.3.2 The *codA*-5-FC method

The expression of the modified *codA* gene under the P_{psbAII} promoter leads to the production of cytosine deaminase. In the presence of 5-FC, the modified cytosine deaminase is able to accept 5-FC as a substrate and catalyses its conversion into 5-FU. The presence of 5-FU leads to the inhibition of protein and DNA synthesis, resulting in cell death. These findings allowed the development of a novel method using *codA*-5-FC for the production of marker-less transformants, with the successful introduction of both *cat* and *LS* genes into the non-coding region between *slr1340* and *sll1255*.

5.3.2.1 Comparison with *sacB*-sucrose method of producing marker-less transformants

Only a single marker-less transformant (containing *cat*) was achieved in all the attempts made using the *sacB*-sucrose method, whereas with the *codA*-5-FC method numerous marker-less transformants were achieved (containing either *cat* or *LS*). When a formal comparison was attempted between the *sacB*-sucrose method and the *codA*-5-FC method, no transformants were obtained using the *sacB*-sucrose method. At least six transformants each containing *cat* and *LS* were obtained using the *codA*-5-FC method. However, as the experiment was not repeated, a fair comparison between the two methods cannot be made. In the literature there are no reported problems with the *sacB*-sucrose method but this may be down to only reporting positive results. One advantage of using the *codA*-5-FC method is that selection is not made on media containing sucrose, which is more susceptible to contamination.

5.3.2.2 Comparison with other methods of producing marker-less transformants

The validation of the *codA*-5-FC method adds to a growing list of methods that can be used to produce marker-less transformants. An advantage of the *codA*-5-FC method is the absence of a scar site seen when recombinases are used to remove antibiotic resistance

genes. The *codA*-5-FC method could be used in strains that cannot use the *sacB*-sucrose method successfully to select for marker-less transformants.

5.3.3 Future work

5.3.3.1 Optimisation of the *codA*-5-FC

Other than the *sacB*-sucrose method, alternative methods of producing marker-less transformants have not been widely used. For the *codA*-5-FC method to become a routine method in cyanobacterial laboratories other improvements could be made to the methodology.

When using the *codA*-5-FC method to introduce *cat*, there was a misalignment with the *codA*/Km^R cassette, which resulted in an extra region of homology that could lead to the production of unwanted gene arrangements. For the introduction of *LS* the misalignment was removed, and only two regions of homology were present; however, there were still transformants when screened that produced bands of unexpected sizes (Figure 5.21).

Extending the flanking regions required for homologous recombination appeared to improve the transformation efficiency as transformants containing *cat* were achieved using *codA*/5-FC, which was not possible previously. There is some evidence to suggest that increasing the region of homology increases the transformation efficiency (Zang et al., 2007). Work done previously in the Purton group demonstrated that when performing natural transformation in *Synechocystis* with a transforming plasmid, transformants could be obtained even when the region of homology upstream and downstream is a little as 30 bp. The work also showed that by increasing the length of the flanking region, the number of transformant colonies recovered was increased and there was a 10-fold increase when the region of homology was increased from 400 bp to 1000 bp (Al-Haj, 2014). Although the selection method employed in this chapter differed, the result suggests that increasing the flanking region may have played a role in obtaining successful transformants. A search of the literature suggests that flanking regions of at least 500 bp upstream and downstream are normally used with the *sacB*-sucrose method (Lagarde et al., 2000; Lea-Smith et al., 2013, 2014).

When performing a transformation, at least three successful transformant lines are required to perform triplicate studies. Screening large number of initial transformants is relatively quick and easy by comparing the growth 'spot' test of transformants on BG-11 and BG-11 containing kanamycin. This helps circumvent the main issue with the *codA*-5-FC method – the high frequency of false positives on the control plate where no DNA has been added. Indeed, there were occasions where more colonies were present on the

control plate than on the transformed plate. When this occurred previously (when transforming the plasmid with pIR706.cat) the plates were dismissed as unsuccessful but this might not have been the case.

Although marker-less transformants produced using *codA*-5-FC is possible using the 5-FC method (Figure 5.28), future work should focus on optimising the conditions (i.e. recovery time before selection pressure is applied) to ensure marker-less transformants are consistently produced, and the number of false positives is reduced to a minimum.

The method should also be attempted in other strains of cyanobacteria, to see whether the technique is transferrable.

5.3.3.2 Utilising the *codA*-5-FC method for introducing an operon.

The overall goal of the 'tools development' was to develop a pipeline in which transgenic operons could be introduced into a neutral genomic locus using a marker-less strategy, thereby establishing the methodology for metabolic engineering of *Synechocystis*. In Chapter 6 the *codA*-5FC method is used to try and introduce a biosynthetic pathway for the production of (*S*)-styrene oxide in *Synechocystis*.

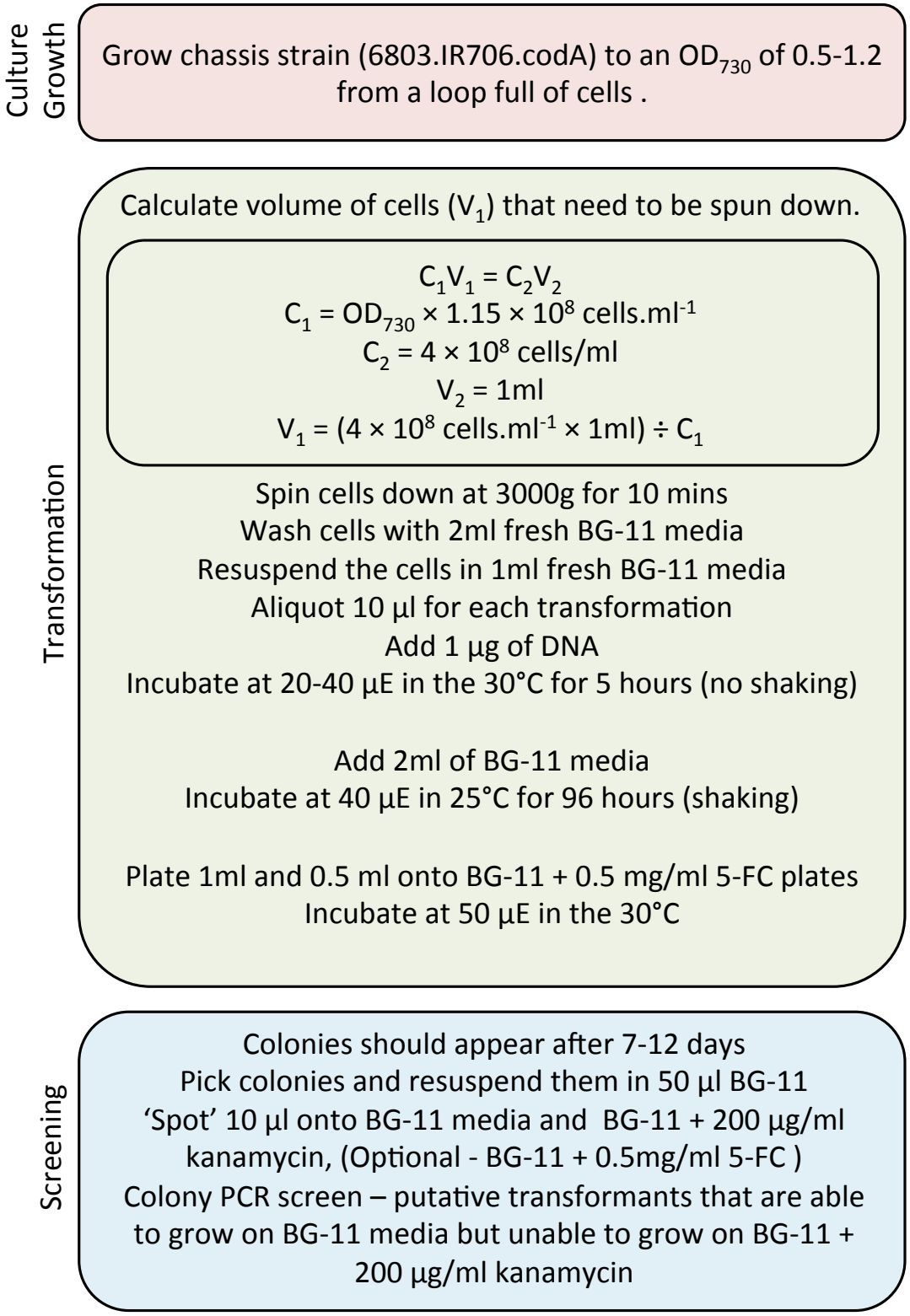


Figure 5.28 Protocol used to produce marker-less transformants in *Synechocystis* using the *codA* – 5-FC method.

Chapter 6 Introducing a heterologous pathway into *Synechocystis* for the production of (S)-styrene oxide

6.1 Introduction

6.1.1 (*S*) – Styrene oxide

(*S*)-Styrene oxide (C_8H_8O) is a chiral, aromatic molecule that is used as a building block for pharmaceuticals, speciality chemicals and cosmetics. The anthelminthic drug, levamisole, and nematocides are made from (*S*)-styrene oxide (McKenna et al., 2013; Mooney et al., 2006). (*S*)-Styrene oxide is traditionally produced from the partial oxidation of styrene over a heavy metal catalyst, however there is no control over the stereochemistry of the product, and further processing is needed to isolate the useful (*S*) enantiomer (McKenna et al., 2013).

The natural production of styrene oxide has been observed in numerous bacteria. Some bacteria in the *Pseudomonas* genus are capable of utilising styrene as a source of carbon (Nishio et al., 2001; Nöthe and Hartmans, 1994; Panke et al., 1998; Wubbolts et al., 1994). Styrene monoxygenase (SMO), encoded by *styAB*, can catalyse the conversion of styrene to styrene oxide, although the enantiomer produced varies from different strains (Otto et al., 2004; Panke et al., 1998). The *styAB* gene from *Pseudomonas* sp. Strain VLB120 and *Pseudomonas putida* S12 encode a SMO capable of producing (S)-styrene oxide in an enantiomeric excess of >99% (Nöthe and Hartmans, 1994; Panke et al., 1998, 2000, 2002).

Both the industrial method and the use of a single enzyme catalyst are limited by the amount of styrene. Styrene is usually produced from the dehydrogenation of ethylbenzene, derived from petroleum, in the presence of steam over a catalyst and the process of production is very energy intensive (Mooney et al., 2006; U.S Department of Energy, 2002).

6.1.2 Engineering styrene and (*S*)-styrene oxide biosynthesis in *E. coli*

In 2011, McKenna and Nielsen engineered *E. coli* to produce styrene from glucose. By utilising the endogenous production of L-phenylalanine (from glucose), styrene was produced by expressing two enzymes, *PAL2* from *Arabidopsis thaliana* and *FDC1* from *Saccharomyces cerevisiae*, in a strain of *E. coli* that over produces L-phenylalanine (Figure 6.1) (McKenna and Nielsen, 2011). PAL2 acts as a phenylalanine ammonia lyase (PAL) and deaminates L-phenylalanine to produce *trans*-cinnamate (tCa). The decarboxylation of tCa to styrene is performed by FDC1, originally labeled as a ferulate decarboxylase, and is the only enzyme shown to carry out this reaction (McKenna and Nielsen, 2011).

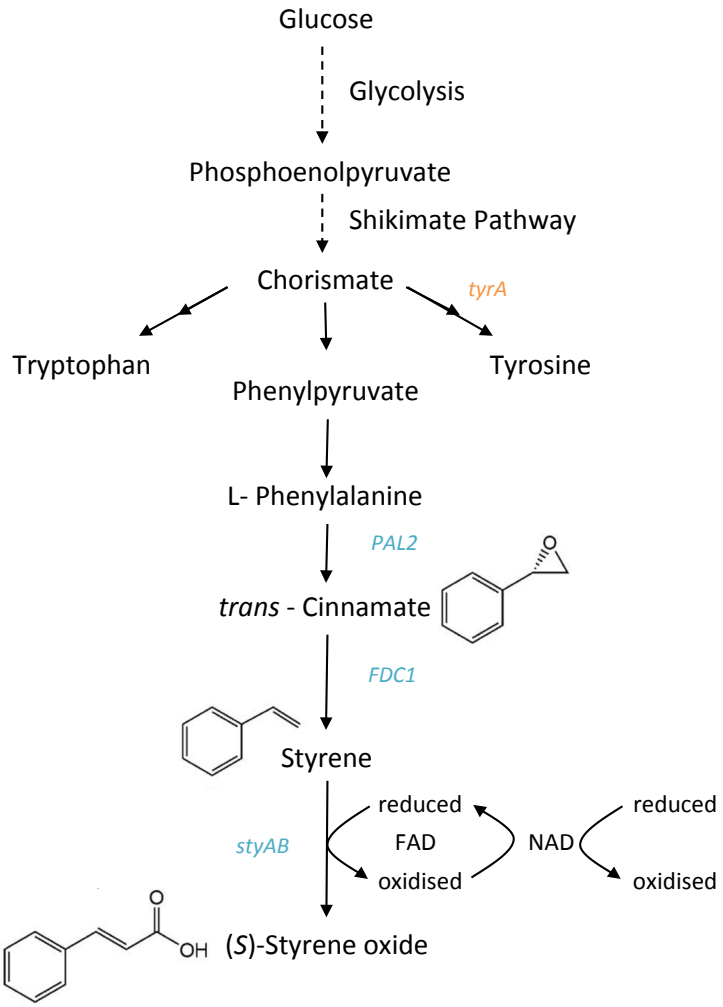


Figure 6.1 Pathway introduced into *E. coli* for the production of (*S*)-styrene oxide.

Modified from (McKenna et al., 2013). Genes introduced into *E. coli* are in blue and *tyrA*, already present, is in orange.

McKenna et al. built upon their work and extended the pathway to produce (*S*)-styrene oxide by expressing StyAB from *Pseudomonas putida* S12, a two component SMO (McKenna et al., 2013). StyA is a FAD-dependent hydroxylase and StyB is a NADH-FAD oxidoreductase, the reaction requires $FADH_2$, which is regenerated through the oxidation of NADH (Otto et al., 2004). The expression of *PAL2, FDC1* and *StyAB* in *E. coli* led to the production of 0.97±0.03 g/L (*S*)-styrene oxide after 72 hours (McKenna et al., 2013). Production was further improved to 1.32±0.03 g/L after 72 hours by increasing the amount of L-phenylalanine available, by deleting *tyrA,* which encodes an enzyme that produces tyrosine from chorismate (also require for phenylalanine production) (McKenna et al., 2013).

The production of (*S*)-styrene oxide has not yet been attempted in *Synechocystis,* but could be a more sustainable method than *E. coli*, as the carbon source required would be produced from photosynthesis.

6.1.3 Aims and objectives

The aim of this project was to demonstrate the use of the molecular tools developed in Chapter 4 and Chapter 5, by introducing a trans-operon (separated by the IGs tested) under the control of P_{cpc560} and P_{nrsB}, using the *codA*-5-FC method into *Synechocystis*. As the production of (*S*)-styrene oxide has not been attempted in *Synechocystis* and has been successfully made in *E. coli* by introducing four genes, the same pathway was chosen to utilise the molecular tools developed so far.

6.2 Results

6.2.1 Toxicity tests

To examine the possible effects of metabolite toxicity that could occur with the introduction of the (*S*)-styrene oxide pathway into *Synechocystis*, tCa, styrene and (*S*)-styrene oxide were added exogenously to cultures of *Synechocystis* to test their effects on growth.

Toxicity tests were performed at an OD_{730} of ~0.1 and ~0.8. Tests starting at OD_{730}: ~0.1 could be seen as representative of expression under the control of P_{cpc560} and at OD_{730}: ~0.8 as representative of expression after nickel-induction using the P_{nrsB} promoter.

When tCa, styrene and (*S*)-styrene oxide was added to a culture of *Synechocystis* with an OD_{730} of 0.8-0.9, the growth of *Synechocystis* appeared unaffected by the addition of tCa up to concentrations of 1 g/L. Styrene appeared to be toxic towards *Synechocystis* at concentrations of 0.2 g/L. After 72 hours, the addition of (*S*)-styrene oxide at 0.5 g/L appeared to have a toxic effect on the cells and after 120 hours the same effect was seen at 0.2 g/L (Figure 6.2).

When a culture of WT *Synechocystis* was grown to an OD_{730} of ~0.1, *Synechocystis* was able to grow in the presence of tCa up to 1 g/L. In the presence of styrene, *Synechocystis* was unable to grow in 0.4 g/L and growth in 0.2 g/L was slower than at lower concentrations of styrene. *Synechocystis* was able to grow in (*S*)-styrene oxide up to 0.5 g/L. In the presence of 1 g/L (*S*)-styrene oxide, the growth of *Synechocystis* appears to be inhibited up until 144 hours after addition (Figure 6.3).

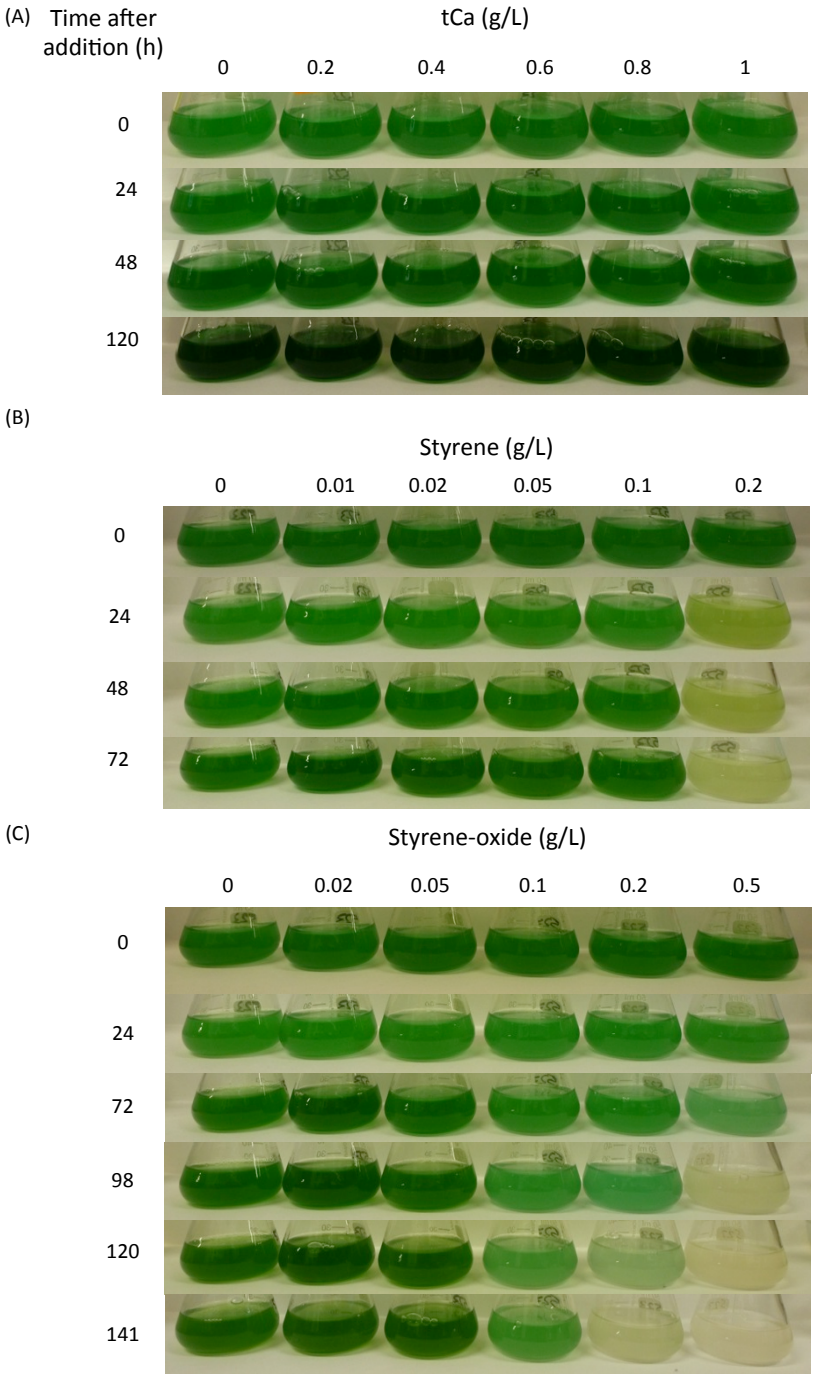


Figure 6.2 Photographs of toxicity tests taken after the addition of tCa, styrene and styrene oxide to a culture of WT *Synechocystis* with a starting OD_{730} of ~0.8

(A) Addition of tCa to cells. (B) Addition of styrene to cells. (C) Addition of styrene oxide to cells.

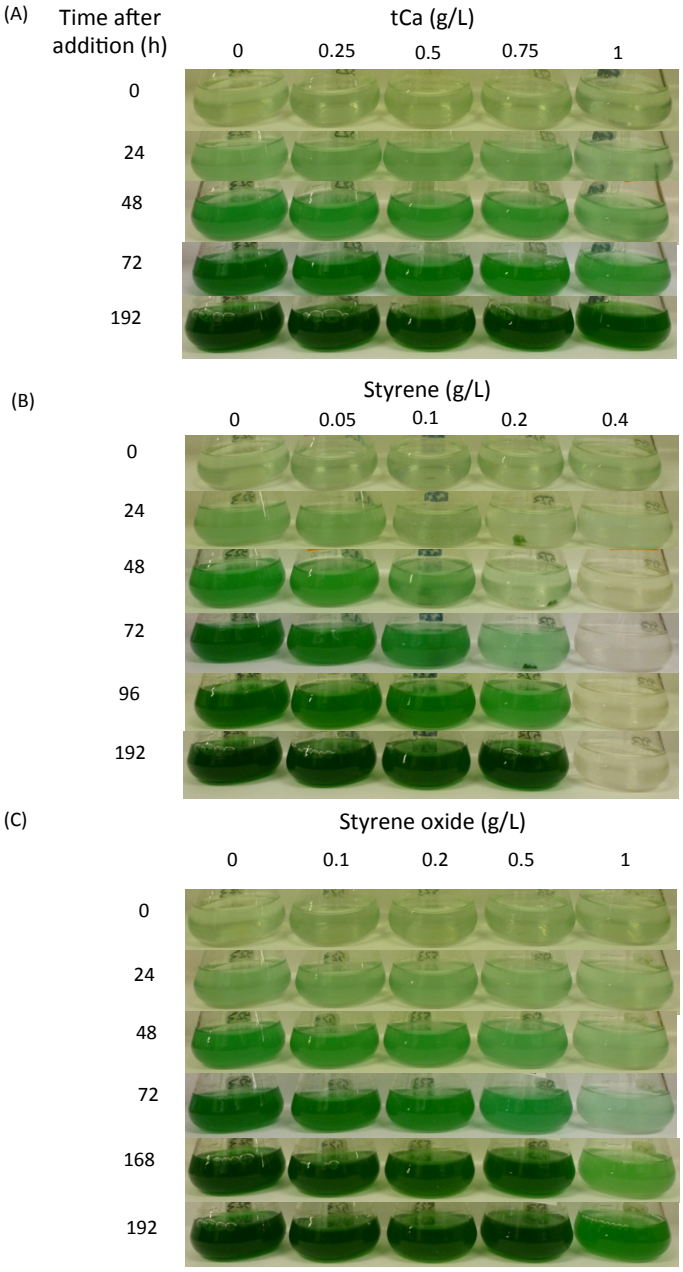


Figure 6.3 Photographs of toxicity tests taken after the addition of tCa, styrene and styrene oxide to a culture of WT *Synechocystis* with a starting OD_{730} of ~0.1

(A) Addition of tCa to cells. (B) Addition of styrene to cells. (C) Addition of styrene oxide to cells.

The results of the toxicity tests suggests that styrene and styrene oxide are toxic towards *Synechocystis*, however the threshold is quite high and *Synechocystis* could produce (*S*)-styrene oxide between 0.1 and 0.2 g/L, which is a high yield for novel compounds produced in *Synechocystis*. If the production of (*S*)-styrene oxide in *Synechocystis* is higher than the toxicity limit then there is the potential to alleviate toxicity by growing the culture in a two-phase system with bis(2-ethylhexyl)phthalate (Panke et al., 2002).

6.2.2 Designing the SO operon

The biosynthetic pathway introduced into *E. coli* led to the successful production of (*S*)-styrene oxide. Therefore, the same genes introduced were chosen for expression into cyanobacteria. To test the molecular techniques developed in the earlier chapters, the genes were designed as an operon, referred to as the styrene oxide (SO) operon, for insertion into *Synechocystis*.

The protein sequence of PAL2 from *Arabidopsis thaliana* (GenBank Accession AY133595), FDC1 from *Saccharomyces cerevisiae* (NP_010828.1) and the StyAB subunits from *Pseudomonas putida* S12 (AJA17113.1 and AJA17114.1) were used to design codon optimised synthesis genes for *Synechocystis* using IDT's Codon Optimization Tool (https://www.idtdna.com/CodonOpt). From the results, codon-optimised sequences that did not contain NdeI and BamHI restriction sites were chosen. An NdeI restriction site was positioned immediately before the start codon for *PAL2* and BamHI was positioned immediately after the stop codon of *StyB*. The functional intergenic regions containing RBS that were investigated in Chapter 4 were used to separate the genes within the operon. The intergenic region IG4 was used to separate *PAL2* and *FDC1* and *StyA* and *StyB* and IG5 was used to separate *FDC1* and *styA* (Figure 6.4).

Due to the large size of the operon (5.5 kb), it was split into two parts for synthesis. An SphI restriction site, approximately halfway in the operon, was selected as the site at which to split the operon into SO_Part1 (2.9 kb) containing *PAL2* and part of *FDC1* gene, and SO_Part2 (2.6 kb), containing the rest of *FDC1* and *styA* and *styB* (Figure 6.4). The SphI site remains in both fragments with a BamHI restriction site immediately after the SphI site in SO_Part1. GeneART synthesised SO_Part1 and SO_Part2 and cloned them into a plasmid containing an ampicillin resistance marker to produce SO_Part1_pMA-T and SO_Part2_pMA-T, respectively (See Appendix).

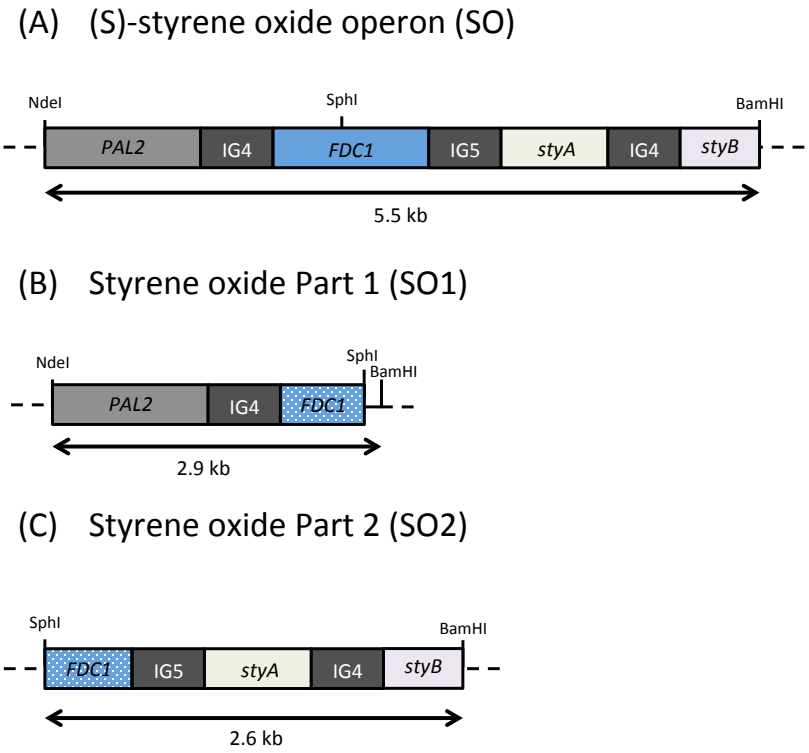


Figure 6.4 Schematic of the SO operon designed for introduction into *Synechocystis*.

The entire (S)-styrene oxide operon (A), was split into two parts: styrene oxide Part 1 (B) and styrene oxide Part 2 (C) for synthesis by GeneART.

6.2.3 Construction of expression plasmids containing the SO operon

The creation of four different transgenic strains of *Synechocystis* containing the styrene oxide operon was planned to study the production of (*S*)-styrene oxide. Two of the four strains would contain the operon at the *psbAII* site, one under the control of the nickel inducible promoter, P_{nrsB} and the other under the P_{cpc560}. To create these two strains the expression vectors, pLAH.nrsB and pLAH.cpc would be used, respectively. The other two strains would express the operon under P_{cpc560} and P_{nrsB}, but at neutral site between *slr1340* and *sll1255*, used previously for producing marker-less transformants using the *codA*-5-FC method. To introduce the genes at this neutral site the pIR706.Ex (see 5.2.4.2.1 for details) expression vector was modified to contain either the P_{cpc560} and P_{nrsB} and a cloning site for the SO operon.

6.2.3.1 Construction of pIR706.Ex.cpc and pIR706.Ex.nrsB

In order to introduce the SO operon into the neutral site between *slr1340* and *sll1255* under the control of the P_{cpc560} and P_{nrsB}, the pIR706.Ex expression vector needed to be modified to introduce the promoters and a cloning site for the SO operon.

The *nrsB* promoter along with the NdeI and BamHI restriction sites was amplified from pLAH.nrsB with primers pLAH.nrsB.F and pLAH.pro.R. The PCR product was then cloned into pIR706.Ex at the HpaI site. Similarly, the *cpc560* promoter along with the NdeI and BamHI restriction sites was amplified from pLAH.cpc with primers pLAH.Pcpc.F and pLAH.pro.R and the PCR product cloned into pIR706.Ex at the HpaI site. PCR performed on with IR706F and IR706R and DNA sequencing confirmed the creation of both new cloning vectors: pIR706.Ex.nrsB and pIR706.Ex.cpc. It should be noted that HpaI is a blunt ended restriction site and, DNA sequencing results confirmed the insertion of the two promoters in their respective plasmids were in the different orientation to one another.

6.2.3.2 Joining the SO operon and constructing pIR706.cpc.SO

SO_Part1 (2.9 kb) was isolated from SO_Part1_pMA-T using NdeI and BamHI, as the vector backbone produced a similarly sized band, the backbone was digested with ScaI. SO_Part2 (2.6kb) was isolated from SO_Part2_pMA-T using SphI and BamHI, and again ScaI was used to digest the backbone. SO_Part1 was cloned into both pIR706.Ex.nrsB and pIR706.Ex.cpc at the NdeI and BamHI sites. Construction of pIR706.Ex.nrsB.SO1 and pIR706.Ex.cpc.SO1 was confirmed by digesting with NdeI and BamHI (Figure 6.5).

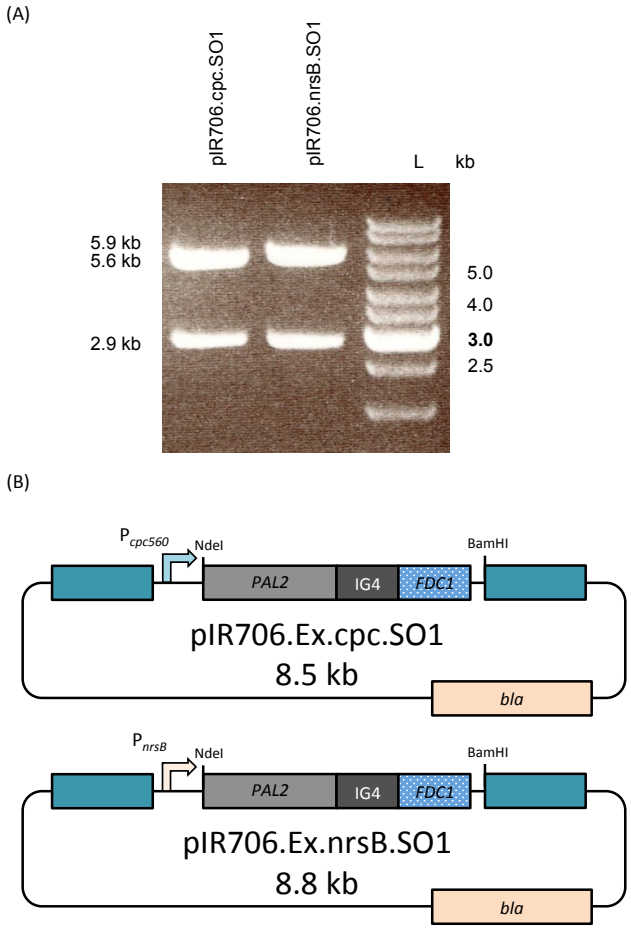


Figure 6.5 Restriction digest to confirm construction of pIR706.Ex.cpc.SO1 and pIR706.Ex.nrsB.SO1

(A) Expected sizes for the pIR706.Ex.cpc.SO1 digest with NdeI and BamHI were 2.9 kb and 5.6 kb. Expected sizes for the pIR706.Ex.nrsB.SO1 digest with NdeI and BamHI were 2.9 kb and 5.9 kb. The sizes are correct for both plasmids. (B) Schematic of expected plasmids pIR706.Ex.cpc.SO1 and pIR706.Ex.nrsB.SO1 with restriction sites of NdeI and BamHI.

SO_Part2 (2.6 kb) was cloned between the SphI and BamHI sites of both pIR706.Ex.nrsB.SO1 and pIR706.Ex.cpc.SO1 to produce pIR706.Ex.nrsB.SO and pIR706.Ex.cpc.SO. It should be noted that *E. coli* colonies carrying putative transformants for pIR706.Ex.nrsB.SO appeared blue and pIR706.Ex.cpc.SO appeared pink, when grown in liquid media; this observation is further discussed at the end of this chapter.

A digest was performed with BamHI, NdeI and ScaI to confirm the construction of both pIR706.Ex.nrsB.SO and pIR706.Ex.cpc.SO, however the digest indicated a problem with pIR706.Ex.nrsB.SO between the BamHI and ScaI restriction site (Figure 6.6). This was confirmed by sequencing analysis using primers nrsBseqF, SO_seq_1R, SO_seq_2F, SO_seq_3R, SO_seq_4F, SO_seq_5F, SO_seq_6R, and SO_seq_EndF. Sequencing analysis also

confirmed the construction of the SO operon in both constructs. Due to time constraints the pIR706.Ex.nrsB.SO was not recreated.

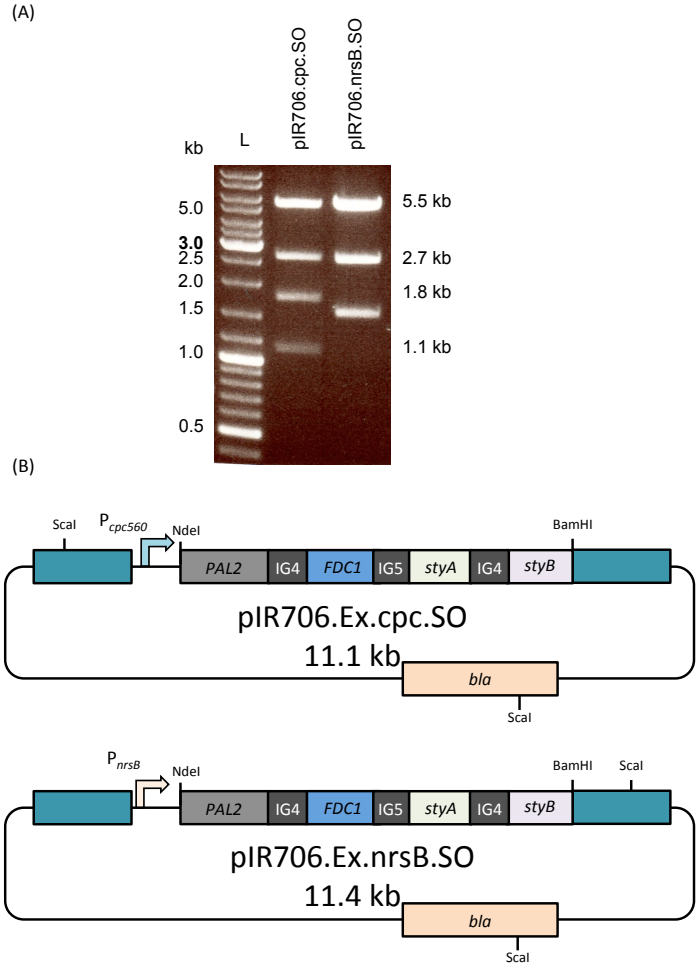


Figure 6.6 Restriction digest to confirm the construction of pIR706.Ex.cpc.SO and pIR706.Ex.nrsB.SO.

(A) Expected sizes for pIR706.Ex.cpc.SO digest with NdeI, BamHI and ScaI were 5.5, 2.7, 1.8 and 1.1 kb. The sizes observed on the gel confirm that this plasmid is as expected. Expected sizes for pIR706.Ex.nrsB.SO were 5.5, 2.7, 2.6 and 0.5 kb. The band sizes observed do not correspond to this with the 0.5 kb band absent and a band at 1.5 kb present instead. (B) Schematic of expected expression plasmids pIR706.Ex.cpc.SO and pIR706.Ex.nrsB.SO with restriction sites of NdeI, BamHI and ScaI.

6.2.3.3 Construction of pLAH.nrsB.SO and pLAH.cpc.SO

The SO operon was digested from the incorrectly constructed pIR706.Ex.nrsB.SO, using NdeI and BamHI and cloned into pLAH.nrsB and pLAH.cpc at the same restriction sites. Putative *E. coli* transformants containing pLAH.nrsB.SO and pLAH.cpc.SO were grown in liquid media with some appearing to produce a blue pigment, as seen previously in *E. coli* strains containing pIR706.Ex.nrsB.SO.

The plasmid extracted from two putative transformants of pLAH.cpc.SO – P1 (normal culture colour) and P4 (Blue pigment present in medium) – were digested with XbaI, BamHI and NdeI together with a putative pLAH.nrsB.SO transformant – N5 (Blue pigment present in medium). The results confirmed the insertion of the SO operon in pLAH.cpc.SO_4 and pLAH.nrsB.SO_5, both of which gave a blue coloration to the culture, whereas for pLAH.cpc.SO_1 the SO operon was absent (Figure 6.7). The production of indigo from these cell lines is the most likely explanation for this result and is further discussed at the end of this chapter.

6.2.4 Creation of transgenic strains 6803.AII.cpc.SO and 6803.AII.nrsB.SO

WT *Synechocystis* was transformed with pLAH.nrsB.SO and pLAH.cpc.SO in order to introduce the SO operon at the *psbAII* site. Colonies appeared after selection for kanamycin resistance. A PCR performed on selected transformants, using primers Ben.seq.up.F and C.frag.BR confirmed the integration of the SO operon in 6 of the 8 tested transformants lines of 6803.AII.cpc.SO (Figure 6.8). Integration of SO operon was also seen in 3 of the 8 putative transformants of 6803.AII.nrsB.SO at the *psbAII* site. All three transformants appeared to be homoplasmic.

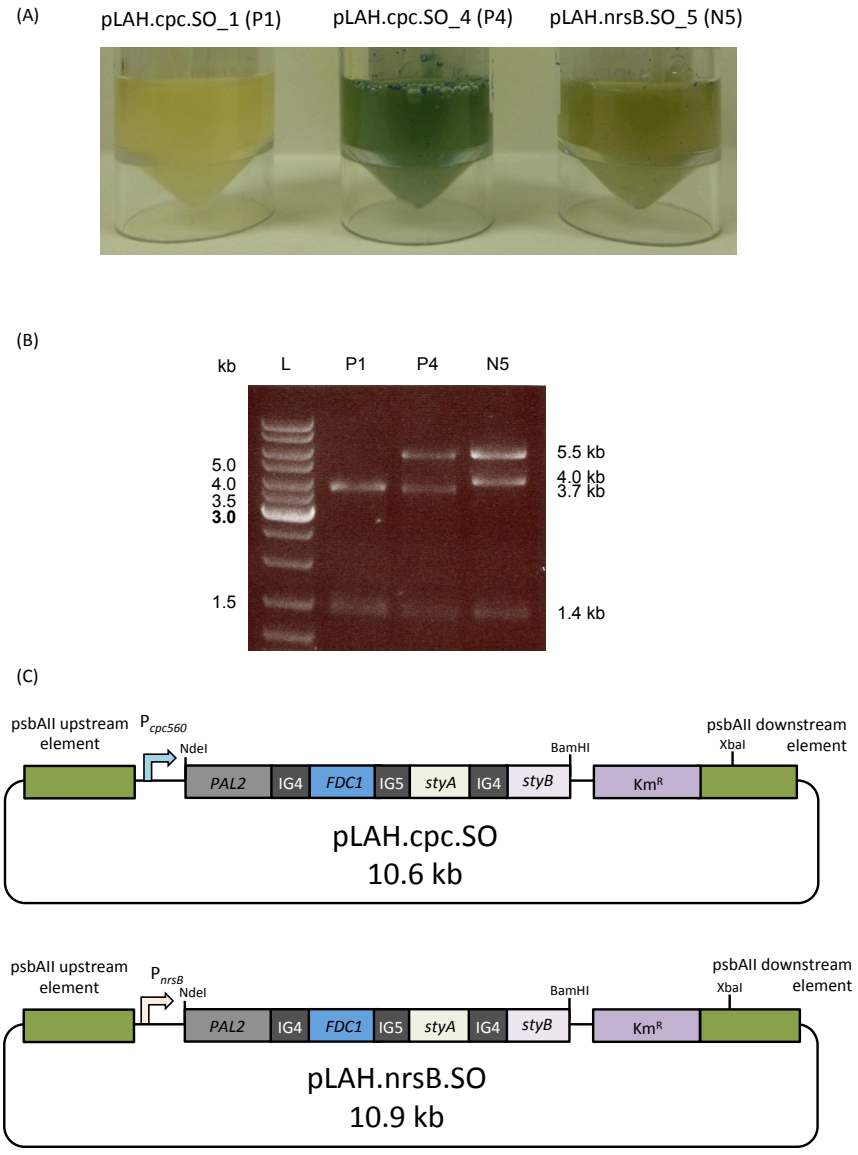


Figure 6.7 Confirmation of the construction of plasmid pLAH.cpc.SO and pLAH.nrsB.SO.

(A) Photograph of selected *E. coli* cultures containing different plasmid constructs. That for pLAH.cpc.SO_1 appears typical of an *E. coli* culture but in both pLAH.cpc.SO_4 (P4) and pLAH.nrsB.SO_5 (N5) a blue pigment appears to be produced. (B) Restriction digest of the plasmids. Expected sizes for pLAH.cpc.SO were 5.5, 3.7, and 1.4 kb. Expected sizes for pLAH.nrsB.SO were 5.5, 4, and 1.4 kb. The expected sizes were seen in the 'blue' cultures, P4 and N5, but not P1. (C) Schematic of pLAH.cpc.SO and pLAH.nrsB.SO with restriction sites of NdeI, BamHI and XbaI.

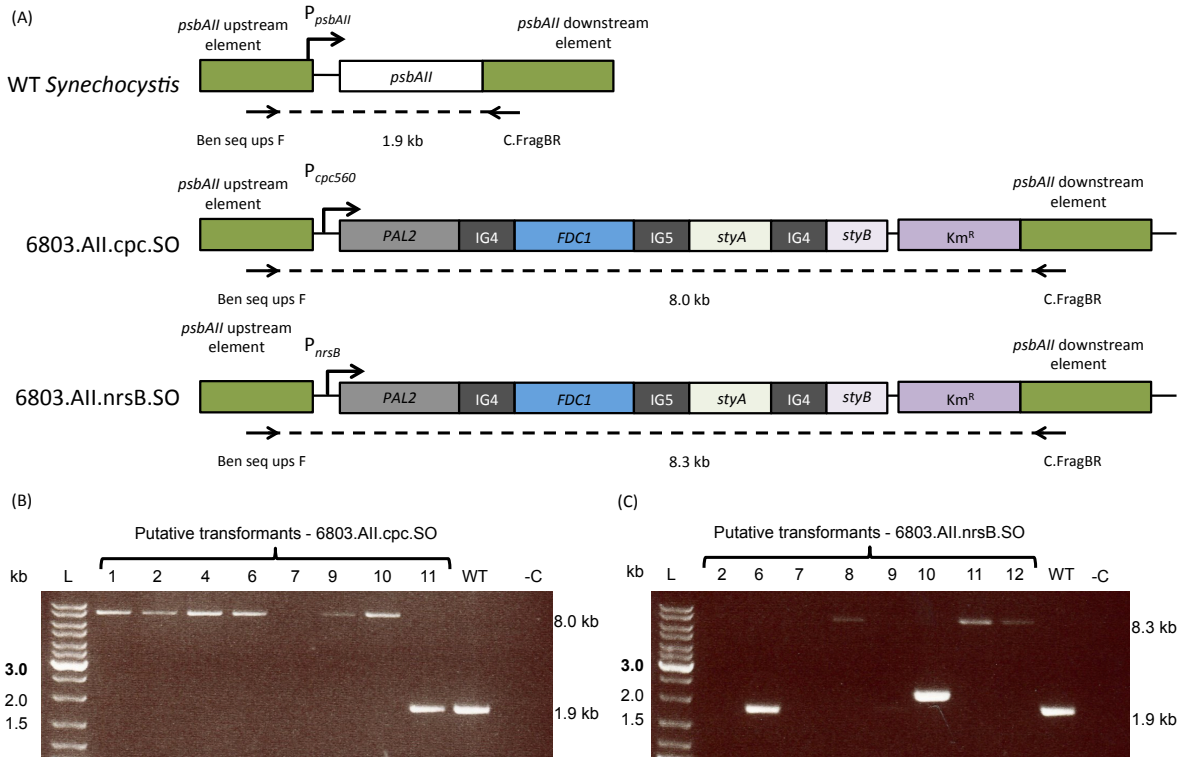


Figure 6.8 PCR screening for successful transformants in *Synechocystis* containing the SO operon at the psbAII locus.

(A) Schematic showing the binding sites of the primers and expected fragment sizes for WT *Synechocystis*, 6803.cpc.SO and 6803.nrsB.SO. (B) PCR screening of putative transformants of 6803.cpc.SO. (C) PCR screening or putative transformants of 6803.nrsB.SO. For (B) and (C) the expected sizes for WT *Synechocystis* was 1.9 kb, 6803.cpc.SO was 8.0 kb and 6803.nrsB.SO was 8.3 kb. WT *Synechocystis* (WT), pLAH.AII.codA (+C) and no DNA template (-C) were used as controls. All successful transformants are homoplasmic with an absent band at 1.9 kb.

6.2.5 Creation of control strain, 6803.IR706.cpc, and unsuccessful attempts at producing marker-less transformants containing the SO operon

Attempts using the *codA*-5-FC method, established in Chapter 5, to produce marker-less transformants with the SO operon using pIR706.Ex.cpc.SO to transform the recipient strain 6803.IR706.codA proved unsuccessful. A transformation using pIR706.Ex.cpc.SO and pIR706.Ex.cpc (to produce a control strain) resulted in the appearance of colonies 13 days after plating. From both sets of plates, 32 colonies were picked and grown onto BG-11 and BG-11 supplemented with kanamycin. Of the 32 putative transformants of 6803.IR706.cpc.SO, all 32 colonies were able to grow in the presence of kanamycin. This indicates that the *codA/Km*^{R} gene cassette was not replaced by the SO operon (results not shown). However, the creation of the control strain, 6803.IR706.cpc was successful. From the 32 colonies picked, five colonies were unable to grow onto BG-11 supplemented with kanamycin but were able to grow on BG-11 only (Figure 6.9). PCR analysis of these five transformant lines confirmed the replacement of *codA* with the P_{cpc560} (Figure 6.9).

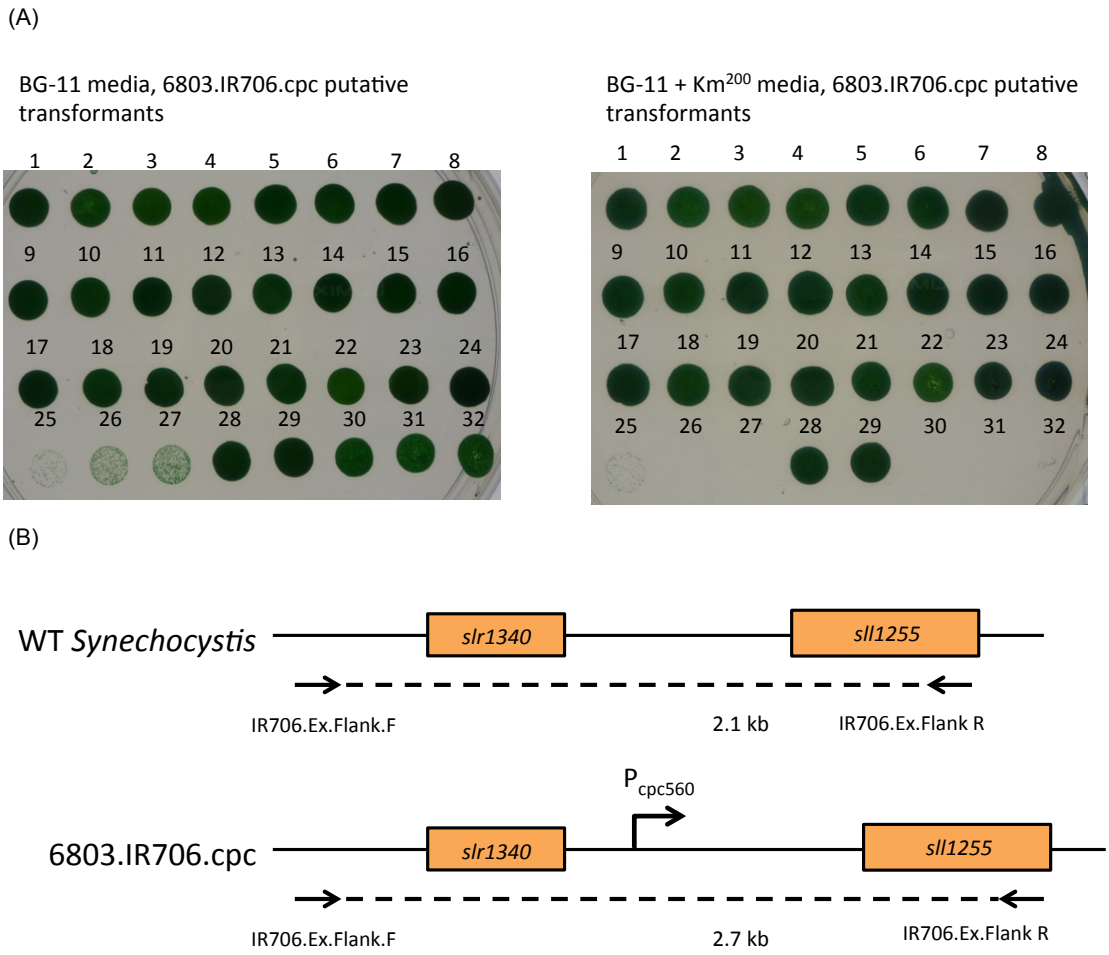


Figure 6.9 Screening of marker-less transformants of *Synechocystis* containing the P_{cpc560} at the neutral site, between slr1340 and sll1255.

(A) Growth 'spot' test of 32 putative marker-less transformants grown on BG-11 and BG-11 supplemented with 200 µg/ml kanamycin. Photograph taken 2 weeks after incubation. From the 32 colonies, 5 appear to be marker-less transformants containing the cpc560 promoter. (B) Schematic showing the primer binding site and expected fragments of WT *Synechocystis* and 6803.IR706.cpc. (C) PCR Screening primers with IR706 Ex Flank F and IR706 Ex Flank R of the five putative transformants. All of those tested contain the cpc560 promoter and are homoplasmic, with a single band seen at 2.7 kb. WT *Synechocystis* (WT) and a no DNA template (-C) are used as controls.

6.3 Discussion

In summary: i) toxicity tests were conducted for (*S*)-styrene oxide and the two intermediates in its synthesis; ii) a synthetic operon for the production of (*S*)-styrene oxide was designed and introduced into *Synechocystis* at the *psbAII* site under the control of the *nrsB* and *cpc560* promoter. However production of (*S*)-styrene oxide has yet to be tested to validate the design of the operon. iii) Introduction of the SO operon into the neutral site using the *codA*-5-FC method was unsuccessful, suggesting more work needs to be done to optimise the selection method when trying to introduce large regions of foreign DNA such as the SO operon.

6.3.1 Toxicity tests

Exogenous addition of tCA, styrene and (*S*)-styrene oxide to WT *Synechocystis* may not cause the same cell response seen if the compound was produced within the cell, but gives an indication of their toxicity towards the cell. The results suggest that *Synechocystis* is able to grow in the presence of tCa up to 1 mg/L. The results differ slightly for styrene and styrene oxide when added to the cells at a low OD and at a higher OD. At a low OD, *Synechocystis* was able to grow in the presence of 0.2 g/L styrene and 1 g/L (*S*)-styrene oxide, but at a higher OD value styrene and (*S*)-styrene oxide are both toxic at 0.2 g/ml. In *E. coli*, the toxic limits were 0.8 g/L for tCa; 0.3 g/L for styrene and 1.6 g/L for (*S*)-styrene oxide (McKenna and Nielsen, 2011; McKenna et al., 2013). So compared to *E. coli*, *Synechocystis* can tolerate a higher concentration of tCa, but a lower concentration of (*S*)-styrene oxide. 1.32 g/L of (*S*)-styrene oxide was produced in *E. coli* after the introduction of the SO operon, which is above the toxicity limit seen in *Synechocystis* (McKenna et al., 2013). In *E. coli* strains where only the SMO is expressed and styrene is provided as a substrate, 12 g/L of (*S*)-styrene oxide was produced when a two-phase system was adopted with bis(2-ethylhexyl)phthalate (Panke et al., 2002). The addition of bis(2-ethylhexyl)phthalate also increased the production of styrene, to 0.8 g/L, in *E. coli* strains expressing *PAL2* and *FDC1* (McKenna et al., 2014). A two-phase system could also be used in *Synechocystis* to alleviate the toxicity of (*S*)-styrene oxide and improve production, however the yields produced in *Synechocystis* may not be capable of competing with *E. coli*.

In *E. coli* there is a correlation between toxicity of the solvent and its log$K_{O/W}$ value, which is not seen in *Synechocystis* (McKenna and Nielsen, 2011; McKenna et al., 2013). The octanol-water partition coefficient (reported as log$K_{O/W}$) of organic solvents often correlates with the antimicrobial activity of the solvent therefore it is used as an indicator of solvent toxicity towards bacteria. The log$K_{O/W}$ of (*S*)-styrene oxide (1.61) is less than

that of styrene (3.05) and tCa (2.13). Usually if the log$K_{O/W}$ calculated is between 1.5 and 5, the solvent is hydrophobic and likely to be toxic (Heipieper et al., 1994; Ramos et al., 2002). Hydrophobic aromatics are thought to accumulate within the cytoplasmic membrane, disrupting the integrity of the membrane, causing ions to leak, a change in pH and the proton gradient (Knoshaug and Zhang, 2008). The toxicity of tCa towards *Synechocystis* should be tested further, as the concentration tested in these experiments did not cause cell death and may not be interacting with the cells to have an effect when added exogenously.

When (*S*)-styrene oxide production was detected in *E. coli*, the intermediates tCa and styrene were not detected, suggesting that there is a rapid turnover of the products (McKenna et al., 2013). If this is the case in *Synechocystis* then the toxicity of (*S*)-styrene oxide is most likely to have a toxic effect towards the cell, compare to styrene and tCa.

6.3.2 Production of indigo

The blue pigment produced in *E. coli* containing plasmids with the SO operon, pLAH.nrsB.SO, pLAH.cpc.SO, pIR706.nrsB.SO and pIR706.cpc.SO was most likely indigo, a brilliant blue pigment used to dye fabrics such as cotton and wool, and was due to the expression of *styAB*. The conversion of indole into indigo in *E. coli* by monooxygenases and dioxygenases, such as the SMO in the SO operon has been demonstrated and are presented in Figure 6.10. *E. coli* transformants expressing *styAB* can therefore be selected for on LB medium supplemented with indole as blue colonies, indicating the production of indigo (Panke et al., 1998). The absence of indole in the medium, may explain why the *E. coli* colonies containing the SO constructs were unable to produce sufficient indigo to be initially seen on solid medium.

The results suggests that *styA* and *styB* are transcribed, although it is not known whether the *Synechocystis* promoters, P_{cpc560} and P_{nrsB} are functional in *E. coli*, or whether there is a cryptic promoter in the upstream flanking region of *styAB* within the SO operon, which is responsible. RT-PCR could be used to check which promoter is transcribing *styA* and *styB*. The amount of indigo produced in the *E. coli* strain carrying the pLAH.cpc.SO, appears to be greater than pLAH.nrsB.SO, which correlates with the higher activity of the P_{cpc560} promoter in *Synechocystis* (Figure 3.22).

On one occasion, transformants containing pIR706.cpc.SO, appeared pink in colour, which could be the production of indirubin (also known as indigo red). One of the proposed mechanisms for the production of indirubin is presented in Figure 6.10. Indirubin can also be produced from the condensation reaction of isatin and indoxyl. Isatin can be made from

oxidation of indoxyl and 2-hydroxylindole (Meyer et al., 2002; Rui et al., 2004). The production of indirubin has been detected when expressing heterologous oxygenases in *E. coli* (Rui et al., 2004).

The production of indigo has not been reported in *Synechocystis* and an assay could be devised to detect the production of indigo by supplementing the media with indole. The blue-green colouration of the cells might make indigo detection in the cultures difficult, however in *E. coli* the insolubility of indigo in the media may only require visual confirmation. The production of indigo would confirm the function of *styAB* in the SO operon.

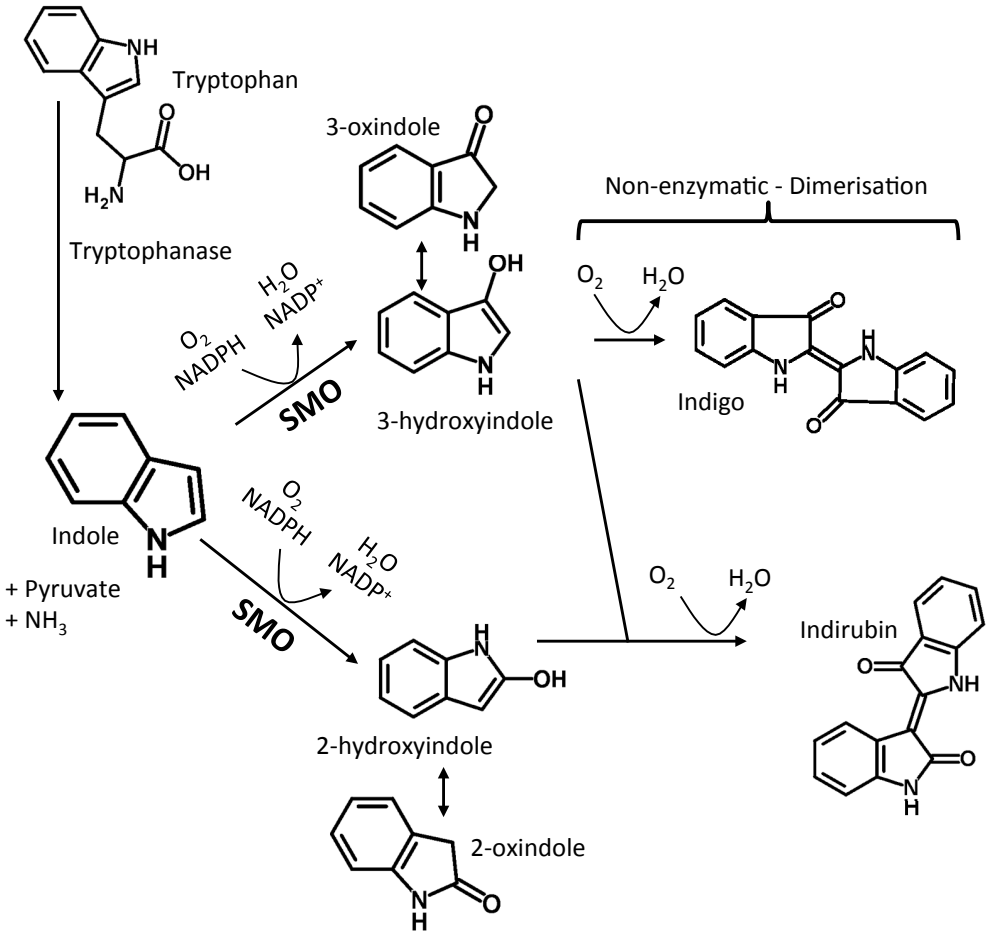


Figure 6.10 Possible pathway for the production of indigo and indirubin in recombinant *E. coli* expressing styrene monoxygenase (SMO).

3-Hydroxyindole is also known as indoxyl. Modified from (Han et al., 2013).

6.3.3 Marker-less transformants containing the SO operon

The few attempts undertaken to introduce the SO operon using the *codA*-5-FC method were unsuccessful, although the creation of the control strain 6803.IR706.cpc, was achieved. This possibly reflects the inefficiency of the method when trying to introduce

large sections of foreign DNA. Furthermore, the protocol used in the previous methods for the successful production of marker-less transformants used 1 µg of plasmid, however as the size of the plasmids containing the SO operon are much larger (more than twice the size) more DNA may be needed to increase the likelihood of obtaining transformants. Finally, the SO operon was cloned into the HpaI site of pIR706.Ex, as with *cat* in Chapter 5, resulting in three regions of homology and the production of transformants that reverted back to the WT strain (see Figure 5.18). Further work is therefore needed to optimise the transformation efficiency of the method and to reduce the number of false positives seen.

6.3.4 Future work

6.3.4.1 Detecting the production of (*S*)-styrene oxide

Due to time constraints, the production of (*S*)-styrene oxide in the transformant lines has yet to be determined. Production of (*S*)-styrene oxide in *E. coli* was detected in the supernatant via HPLC analysis and could be similarly employed to detect the production in *Synechocystis*. If (*S*)-styrene oxide or the intermediates are not detected then RT-PCR could be used initially to detect whether the genes in the operon are being transcribed. The expression of each gene would then need to be confirmed by Western blot analysis. In order to do this, expression tags will need to be redesigned into the construct and retransformed into *Synechocystis*.

6.3.4.2 Optimising the production of (*S*)-styrene oxide

Once (*S*)-styrene oxide has been detected in *Synechocystis*, production could be increased by further metabolic engineering to reduce the effect of any bottlenecks present in the metabolic pathway. In *E. coli* the amount of (*S*)-styrene oxide produced was 0.97±0.03 g/L after 72 hours. However, this was increased to 1.32±0.03 g/L after 72 hours, by deleting *tyrA* which led to an increase in the amount of chorismate available, which in turn increased the pool of phenylalanine available for *PAL2* (McKenna et al., 2013). Attempts were also made in *E. coli* to increase the pool of phenylalanine by deleting *trpE* (involved in the synthesis of tryptophan) along with *tyrA* however the increase of phenylalanine seen with the deletion of *tyrA* was negated by the deletion of *trpE* (McKenna et al., 2013). A similar approach could be done in *Synechocystis* to improve the amount of phenylalanine present for PAL2.

Chapter 7 Final discussion

7.1 Summary of findings

As energy demands increase, photosynthetic microorganisms such as cyanobacteria have the potential to provide a 'green' alternative for the production of biofuels and other chemicals that are usually derived from finite fossil fuels. However, the levels of novel compounds produced in engineered strains are a long way from industrial production levels and yields are not yet economically viable. A significant limitation with engineering cyanobacteria for the production of novel compounds is the lack of a suite of biological parts that are well characterised, and allow advanced, metabolic manipulation.

The initial aims of the project were to investigate metabolic engineering in *Synechocystis* using limonene as a test product that has commercial value as a fragrance and could be a potential replacement for jet fuel, to develop and characterise new DNA parts, a new genomic insertion site and a marker-free transformation strategy for *Synechocystis* and to use these tools to demonstrate multigenic engineering by introducing a synthetic operon for the production of a novel compound.

A synthetic LS gene was successfully expressed in *Synechocystis* under the control of the three promoters (P_{psbAII}, P_{nrsB} and P_{cpc560}), however limonene production was not detected in this study and further work would be needed to understand why this was the case. During the course of the project, limonene production in different cyanobacterial species was achieved by three separate groups (Davies et al., 2014; Halfmann et al., 2014; Kiyota et al., 2014). Expression of *Mentha spicata* LS in *Synechocystis* 7002 using a dodecane overlay, similar to the system attempted in this study, resulted in the highest production of limonene (4 mg/L) out of the three strains (Davies et al., 2014). In the other two strains, different LS genes and different methods of capturing limonene were used, which resulted in lesser yields, even though some genes involved in the MEP pathway were overexpressed to improve the amount of substrate for LS (Halfmann et al., 2014; Kiyota et al., 2014).

The newly discovered 'super'-strong promoter, P_{cpc560}, was shown to be a stronger promoter than P_{psbAII}, and P_{nrsB}, and was used to express multiple genes in an operon. Several IGs regions were tested within an operon and gave varying levels of expression of the downstream gene. Furthermore, it was found that the absence of an IG region still resulted in the translation of the downstream gene.

An alternative method to the *sacB*-sucrose method for the production of marker-less transformants was developed. Although the method requires some optimisation, the production of marker-less transformants enables numerous rounds of transformations

unlimited by the number of available antibiotic resistance genes. The method could be used with other strains of cyanobacteria such as *Synechococcus* 7942 that are able to insert genes into their genome via homologous recombination.

An operon was designed to produce (*S*)-styrene oxide and was successfully introduced into *Synechocystis*, using the IG regions designed in Chapter 4. However, due to time limitations further work is required to detect the production of (*S*)-styrene oxide.

7.2 Challenges and opportunities to using cyanobacteria as a microbial cell factory

The production of novel compounds has been demonstrated in cyanobacteria at a lab-scale, however before these products can be produced at an industrial level there are many challenges ahead. In this section, some of these challenges are discussed.

7.2.1 Improving product yield

The yields of novel compounds produced from engineered *E. coli* or *S. cerevisiae* are significantly greater than those seen in cyanobacteria (Liao et al., 2016). Out of the three cyanobacterial strains that managed to produce limonene the engineered *Synechococcus* 7002 produced the most (4 mg/L) but the yield was still approximately 100 times less than that achieved in *E. coli* (400 mg/L)(Alonso-Gutierrez et al., 2013; Davies et al., 2014). Various strategies have been investigated in the hope of increasing the production levels in cyanobacteria.

7.2.1.1 Engineering the chassis

As photosynthesis is driving the production of novel compounds in engineered cyanobacteria, improvements to the photosynthetic efficiency have the potential to improve yield. Improving light harvesting and CO_2 fixation in cyanobacteria are all strategies that have been investigated in cyanobacteria (Jensen and Leister, 2014).

Expanding the solar spectrum used in photosynthesis from 400-700 nm to 750 nm could be a method of extending the wavelengths absorbed and improve light harvesting. This could be achieved by introducing chlorophyll *d* found in *Acaryochloris marina*, and a recently discovered chlorophyll *f* found in *Halomicronema hongdechloris,* which absorb light in the infra-red region. This could increase the available photons for photosynthesis by 19%, but whether this could help increase biofuel production is unknown (Chen and Blankenship, 2011).

As aforementioned in the introduction, RuBisCO has been the target for many studies to enhance carbon fixation but there has been little research investigating alternative strategies (Durall and Lindblad, 2015). Recently, the overexpression of RuBisCO, and a few genes involved in the Calvin-Bassham-Benson cycle resulted in improved growth rate and biomass accumulation in *Synechocystis* (Liang and Lindblad, 2016). To improve carbon uptake, the endogenous bicarbonate transporters found in *Synechocystis* were overexpressed. This led to an increase in HCO_3^- uptake and resulted in increased biomass accumulation (Kamennaya et al., 2015). These findings should encourage further studies

to see the effect of improved carbon fixation and light harvesting on the production of novel compounds in cyanobacteria.

7.2.1.2 Relieving toxicity

E. coli is capable of producing 400 mg/L limonene. However, if limonene is to be produced as a competitive replacement for jet fuel, it has been calculated that approximately 6.4 g/L of limonene needs to be produced, which is 0.76% (v/v). Since complete growth inhibition is seen in *E. coli* with 0.025% (v/v), a more tolerant strain is needed (Dunlop, 2011). Recently, the oxidised derivative of limonene, limonene hyperoxide, was identified as the cause of the toxic effect seen in *E. coli*, rather than limonene. The toxicity of limonene hyperoxide could be alleviated by a point mutation in an alkyl hyperoxidase, AhpC (Chubukov et al., 2015). The effect of limonene hyperoxide has not been investigated in cyanobacteria and it would be interesting to examine whether a similar effect is seen. If this is the case in *Synechocystis*, *sll1621* is predicted to encode an AhpC/TSA family protein and could be studied further to see whether it could relive toxicity.

Transcriptomic and proteomic analyses performed on *Synechocystis*, in the presence of ethanol, butanol and hexane, have helped to identify genes that are up- or down-regulated. Potentially, these could be targeted in order to improve tolerance (Anfelt et al., 2013; Liu et al., 2012; Qiao et al., 2012; Tian et al., 2013; Wang et al., 2012; Zhu et al., 2013). For both butanol and ethanol treatment of *Synechocystis*, specific response regulators (encoded by *sll0039* and *sll0794*, respectively) have been identified from proteomic data as potential targets (Niu et al., 2015; Song et al., 2014). Butanol tolerance was increased when the general stress response regulator was overexpressed (Kaczmarzyk et al., 2014). Proteomic and transcriptomic data following exposure to other novel compounds should be investigated to identify possible targets that could help relieve toxicity.

Finally, the expression of efflux pumps, a class of membrane transporters that use the proton motive force to export toxic compounds out of the cell, has been demonstrated to improve tolerance to limonene in *E. coli* (Dunlop et al., 2011). The use of efflux pumps to improve tolerance has not been well researched in cyanobacteria but could prove useful, not just in alleviating toxicity, but also by potentially improving yield.

7.2.1.3 Diverting carbon flux

In *Synechocystis* approximately 20 times more carbon goes towards the synthesis of carbohydrates than towards the synthesis of terpenoids (Rosgaard et al., 2012). This may explain why limonene is not produced to such high levels in cyanobacteria in comparison

to *E. coli*. To improve the production of limonene, more fixed carbon needs to be rerouted towards the MEP pathway.

There are a number of examples of modifying competing pathways in order to increase yield. The pathway for the synthesis of 3-hydroxybutyrate, the monomer form of PHB, was introduced in *Synechocystis* and production was increased when the native competing pathway, producing the polymer PHB, was removed (Wang, 2013). From *in silico* metabolic flux analysis, carbon flux through acetyl-CoA appeared to limit the photoautotrophic production of biofuels. The pool of acetyl-CoA available was increased by expressing a phosphoketolase, this resulted in improved production of butanol (Anfelt et al., 2015).

Various metabolic reconstructions have been made for *Synechocystis*, to improve the understanding of the metabolic network present and use this information to improve the production of novel compounds (Erdrich et al., 2014; Knoop and Steuer, 2015; Montagud et al., 2010, 2011; Nogales et al., 2013). One of the first genome-scale models of *Synechocystis* predicted the effect of introducing an ethanol fermentation pathway on growth and production levels which matched experimental data (Fu, 2009). The use of genome scale models can assist a more directed and efficient approach to metabolic engineering. The majority of the models produced have been developed using *Synechocystis*. However, in 2014 the first genome scale model for *Synechococcus* 7942 was produced (Triana et al., 2014) and there have also been attempts to model the nitrogen-fixing genus of cyanobacteria, Cyanothece (Mueller et al., 2013; Saha et al., 2012).

7.2.2 Controlling gene expression

The metabolic engineering of a microorganism such as *Synechocystis* requires the coordinated expression of multiple transgenes. As the ability to tightly control gene expression is important when attempting to redesign the cellular machinery, there has been increasing interest in finding and characterising suitable promoters. Other mechanisms for controlling gene expression have also been investigated.

7.2.2.1 Promoters

E. coli inducible promoters have been modified to improve their efficiency in cyanobacteria, but this has been to varying levels of success (Heidorn et al., 2011; Huang and Lindblad, 2013). A range of promoters were developed with differing levels of expression by altering the distance between the -10 and -35 polymerase sigma binding units (Albers et al., 2015). The nickel inducible promoter can be used to turn on expression of transgenes. However, the presence of low levels of nickel in any medium

means that there is a degree of leaky expression even in the absence of added nickel. Furthermore, the use of heavy metals for induction is not ideal when scaling production (Abe et al., 2014). A green light inducible promoter has been developed and avoids the issues associated with using a chemical inducer (Abe et al., 2014). However, when production is up-scaled it would be limited to growth in photobioreactors, as cultures would need to grow in red light and expression is not completely switched off in red light (Abe et al., 2014).

The discovery of P_{cpc560} has also shown the potential importance of TFBSs in the expression of genes in cyanobacteria (Zhou et al., 2014). The lack of TFBSs in strong *E. coli* promoters has been hypothesised as an explanation as to why these promoters do not perform as well in cyanobacteria (Zhou et al., 2014). In future work, by altering the 14 predicted TFBSs within P_{cpc560}, a range of promoters could be produced with varying strengths of expression. Also, when designing new promoters, native positive TFBSs could be introduced to improve the strength of the promoter.

7.2.2.2 RNA control

Riboswitches are mRNA elements that can be used to regulate post transcriptional steps of gene expression. Riboswitches comprise of an aptamer sequence often located in the 5' untranslated region (5'UTR) of mRNA that is the binding site for a specific ligand. When the ligand is bound to the aptamer sequence this leads to a conformational change that then affects the expression of the adjacent gene (Nakahira et al., 2013). Riboswitches can work to either affect translation initiation or attenuate transcriptional termination. Theophylline-dependent riboswitches have been developed for *Synechococcus* 7942, *Synechocystis*, *Anabaena* sp. strain PCC 7120 and *Leptolyngbya* sp. strain BL090 (Ma et al., 2014; Nakahira et al., 2013; Ohbayashi et al., 2016). One theophylline-dependent riboswitch was very effective at repressing expression in the absence of the ligand and was capable of a 190-fold increase in the presence of the ligand (Nakahira et al., 2013). Another advantage to using the riboswitches is the ability to tune the expression, by altering the concentration of ligand.

Riboregulators have been used in cyanobacteria, and are made up of two components: a *cis*-repressive mRNA (crRNA) and *trans*-activating RNA (taRNA). The crRNA is repressed by its 5' UTR, preventing the ribosome from binding to the RBS. The production of the taRNA, results in the formation of a hybrid complex with the 5' UTR of crRNA, which exposes the RBS site to enable translation. Recently, a scaffold fused-riboregulator from *E.*

coli was successfully adapted and used in *Synechocystis* as a method of controlling gene expression (Sakai et al., 2015).

The role of antisense RNA (asRNA) in cyanobacteria is thought to play an important role in controlling gene expression. Transcribed from the antisense strand of a gene, asRNAs are found either directly antisense to the gene or within 50 bp upstream or downstream (Kopf and Hess, 2015). In *Synechocystis*, one third of transcriptional start sites were on the complementary strand of 866 genes suggesting an important functional role for antisense transcription (Mitschke et al., 2011). In *Synechocystis,* asRNAs may repress or induce expression of genes (Dühring et al., 2006; Sakurai et al., 2012). The presence of asRNA has also been demonstrated in *Anabaena* sp. PCC 7120 (Hernández et al., 2006) and two strains of *Prochlorococcus* (Voigt et al., 2014). AsRNA has been used to repress genes expression in *Synechococcus* 7942 and *Anabaena* 7120 (Higo et al., 2016; Holtman et al., 2005).

7.2.2.3 Ribosome binding sites

From this study, it was demonstrated that within a two-gene operon the presence of a RBS was not required for translation of the downstream gene. In order to develop RBS tools in cyanobacteria further research is needed to understand protein translation initiation and the importance of the SD sequence in cyanobacteria, especially as is it not well understood and most information is based on what is known from *E. coli*. There are tools such as the RBS Calculator available for predicting the effect of a RBS, and such tools are effective at predicting translation in *E. coli.* However, in *Synechococcus* 7002, the predicted levels did not correlate well with the experimental data (Markley et al., 2015).

7.2.3 Scaling up production

Although cyanobacteria have been successfully engineered to produce biofuels and other chemicals at the lab-scale there are still a lot of challenges that need to be overcome before commercialisation (Parmar et al., 2011). There are some technical challenges, regarding cultivation and harvesting, and ecological challenges that need to be considered to ensure the environmental benefits of using cyanobacteria are achieved. There are also economic challenges related to the cost of scaling up production of biofuels (Parmar et al., 2011).

Although cyanobacterial cultivation would not use as much land as crop plants, like plants they still need a source of major nutrients such as phosphorus and nitrogen. There has been limited research using strains that are capable of nitrogen fixing for the production of biofuels, which would circumvent the issues of using nitrogen fertilisers produced using the Haber-Bosch process. Another consideration is that strains such as *Synechocystis* and

Synechococcus 7942 are freshwater strains. However, large-scale cultivation puts significant demands on freshwater leading to competition with drinking water and water used for crop irrigation. Ideally, more research should be done to find strains that are capable of growing in wastewater or salt water (Nozzi et al., 2013).

One of the main problems with scaling cyanobacterial cultivation is light delivery. When scaling up, self-shading limits the depth of the culture, although rapid mixing can help to address this. Furthermore, only 10-15% of light is thought to be available to cells for photosynthesis (Nozzi et al., 2013). The growth of cyanobacteria using an open pond system is significantly cheaper than in closed photobioreactors. However, the amount of biomass produced in an open pond is significantly lower and there are environmental issues over the escape of engineered strains of cyanobacteria (Sarsekeyeva et al. 2015). There is very little research in cyanobacteria into biocontainment mechanisms, to prevent the release of any mutagenic strains into the environment. The expression of a DNA/RNA nuclease NucA under different metal inducible promoters was investigated as a possible suicide switch, although further research and improved inducible promoters are needed to develop this system (Čelešnik et al., 2016).

The genetic and metabolic networks in *E. coli* and *S. cerevisiae* are well characterised and better understood in comparison to those of cyanobacteria. This has allowed the development of a wide range of synthetic biology tools available to carry out advanced and predictable genetic engineering (Scaife et al., 2015). The main advantage cyanobacteria have over *E. coli* and *S. cerevisiae* is the ability to fix CO_2 and in some cases to fix N_2, rather than requiring an exogenous source of fixed carbon such as sugars or organic acids and fixed nitrogen such as nitrates, nitrates, ammonium or urea. It might therefore be possible to use a mixed cultivation approach to couple the metabolic engineering benefits of the *E. coli* or yeast platforms with the autotrophy of cyanobacteria. Indeed, there has been some research investigating the growth of cyanobacteria engineered to secrete sugars required to support the production of novel compounds in *E. coli* (Niederholtmeyer et al., 2010). If such a cyanobacterium was also capable of fixing nitrogen it could also be used to provide a nitrogen source.

7.2.4 The future of cyanobacterial biotechnology

Over the next few years the research areas discussed below have the potential to significantly advance the development of cyanobacterial engineering in a predictable manner.

7.2.4.1 Chassis strain

Cyanobacteria represent such a diverse phylum of species, and only a few strains have been studied in any detail. The diversity in phenotypes and genotypes can be utilised for various applications and the production of different novel compounds. However, one problem with the diversity seen in cyanobacteria is the biological tools developed in one strain may not be transferrable across species without laborious evaluation and modification.

Although research continues to improve the synthetic biology tools available for cyanobacteria, the slower growth rates of model cyanobacteria compared to *E. coli* has an effect on the time it takes to produce transformant lines and develop and characterise different tools. In *E. coli* this process may only take a few days, whereas in *Synechocystis* this can take months. *Synechococcus elongatus* UTEX 2973 (*Synechococcus* 2973) was recently identified, as having a doubling time similar to that of *S. cerevisiae*, at 1.9 hours, which is much faster than *Synechococcus* 7942, which has a doubling rate of 4.9 hours (Yu et al., 2015). DNA can be introduced by conjugation but the strain cannot yet be naturally transformed, limiting its use (Wendt et al., 2016; Yu et al., 2015). Nevertheless, an engineered strain of *Synechococcus* 2973 successfully produced 8.7 g/L of sucrose over 21 days (Song et al., 2016).

So far, *Synechococcus* 7002 has not been utilised to the same extent as the freshwater strains *Synechocystis* or *Synechococcus* 7942. However as a marine strain, *Synechococcus* 7002 is attractive as an industrial platform, as it is capable of growing in high salt concentrations, high light and under a range of high temperatures between 30 and 38°C (Ruffing, 2014). One of the limitations with using this strain is the lack of genetic tools available. However, there have been recent progress with the construction of a few tools that enable the control of gene expression and a method for selecting marker-less transformants (Begemann et al., 2013; Markley et al., 2015; Zess et al., 2016).

7.2.4.2 CRISPR/Cas9 and CRISPRi

Clustered regularly interspaced short palindromic repeats (CRISPR) along with CRISPR associated proteins (Cas proteins) provide an adaptive immune system in bacteria against invading viruses and plasmids. The CRISPR system can be utilised for genome editing by creating double stranding breaks at a specific site, which can be used to create mutants, deletions or introduce DNA (Jinek et al., 2012). The use of CRISPR in cyanobacteria is still in its infancy. The first reported use of the CRISPR/Cas9 system was earlier this year, as was used in *Synechococcus* 2973 to generate marker-less mutants (Wendt et al., 2016).

During this study it was shown that the expression of *cas9* was toxic to the cell, unless it was transiently expressed, which may explain why the system has not been used much in cyanobacteria (Wendt et al., 2016). The use of CRISPR/Cas9 genome editing systems could progress cyanobacteria as a platform further, especially for strains where efficient genome manipulation via homologous recombination is lacking.

CRISPR interference (CRISPRi) is an attractive method to repress gene expression at the level of transcription and has so far been investigated in *Synechocystis* and *Synechococcus* 7002 (Gordon et al., 2016; Yao et al., 2016). In *Synechocystis*, CRISPRi was used to repress the formation of PHB, known to compete with biofuels for acetyl-CoA, and is also capable of repressing multiple genes at the same time (Yao et al., 2016). CRISPRi could therefore be used to divert flux from competing pathways, without abolishing the pathways completely (Gordon et al., 2016).

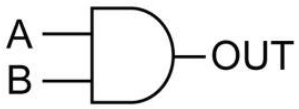
7.2.4.3 Genetic Circuits

By designing and creating genetic circuits within cyanobacteria, cells can be programmed to respond to a greater number of inputs (pH, temperature, light etc.) that can then be used to increase biological control (Wang et al., 2011). Synthetic gene-regulatory circuits have been constructed in *E. coli* to perform logic functions, with the AND and NOR gates (Figure 7.1) tested in *E. coli* under industrially relevant conditions that could be used for fermentation (Gardner et al., 2000; Moser et al., 2012). Recently, a two-input AND gate was developed in *Synechocystis*, responsive to low oxygen and high anhydrotetracycline (Immethun et al., 2016). The discovery of more promoters and riboswitches will enable the creation of more genetic circuits using different logic gates (i.e. OR, NOT, XOR, NAND) (Figure 7.1) that could provide tighter control of the synthetic biological circuits.

7.2.5 Final thoughts

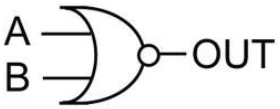
The potential of utilising cyanobacteria's ability to use sunlight to fix carbon for the production of novel compounds in engineered strains is significant. There are a lot of challenges ahead regarding scaling up however the most pressing concern over the next few years for cyanobacteria seems to be ensuring metabolic pathways can be introduced and controlled in a predictable manner. By developing new and improved biological parts and molecular techniques that are transferrable between strains, finding new strains and building detailed metabolic models; the true capabilities of cyanobacteria can be explored. Cyanobacteria can then be evaluated to see what novel compounds are best suited for production in these hosts over other model microorganisms.

A) **AND** Gate



Input		Out
A	B	
0	0	0
1	0	0
0	1	0
1	1	1

B) **NOR** Gate



Input		Out
A	B	
0	0	1
1	0	0
0	1	0
1	1	0

C) Other Logic Gates

NAND

Input		Out
A	B	
0	0	1
1	0	1
0	1	1
1	1	0

XOR

Input		Out
A	B	
0	0	0
1	0	1
0	1	1
1	1	0

OR

Input		Out
A	B	
0	0	0
1	0	1
0	1	1
1	1	1

NOT

Input A	Out
0	1
1	0

Figure 7.1 Truth tables of logic gates used in genetic circuits.

A) The AND gate only gives an output, when both inputs represented by A and B are present. B) The NOR gate only gives an output when both inputs are absent. C) Truth tables of NAND, XOR, OR and NOT gates that could be utilised in a genetic circuit.

References

Abe, K., Miyake, K., Nakamura, M., Kojima, K., Ferri, S., Ikebukuro, K., and Sode, K. (2014). Engineering of a green-light inducible gene expression system in *Synechocystis* sp. PCC6803. Microb. Biotechnol. *7*, 177–183.

Aikawa, S., Nishida, A., Ho, S.-H., Chang, J.-S., Hasunuma, T., and Kondo, A. (2014). Glycogen production for biofuels by the euryhaline cyanobacteria *Synechococcus* sp. strain PCC 7002 from an oceanic environment. Biotechnol. Biofuels *7*, 88.

Aikens, J., and Turner, R.J. (2013). Transgenic photosynthetic microorganisms and photobioreactor.

Akiyama, H., Okuhata, H., Onizuka, T., Kanai, S., Hirano, M., Tanaka, S., Sasaki, K., and Miyasaka, H. (2011). Antibiotics-free stable polyhydroxyalkanoate (PHA) production from carbon dioxide by recombinant cyanobacteria. Bioresour. Technol. *102*, 11039–11042.

Albers, S.C., Gallegos, V.A., and Peebles, C.A.M. (2015). Engineering of genetic control tools in *Synechocystis* sp. PCC 6803 using rational design techniques. J. Biotechnol. *216*, 36–46.

Al-Haj, L.A. (2014). Development of genetic engineering tools for the cyanobacterium *Synechocystis* PCC 6803 for advanced biofuel production. Ph.D. University College London (University of London).

Alm, E.J., Huang, K.H., Price, M.N., Koche, R.P., Keller, K., Dubchak, I.L., and Arkin, A.P. (2005). The MicrobesOnline web site for comparative genomics. Genome Res. *15*, 1015–1022.

Alonso-Gutierrez, J., Chan, R., Batth, T.S., Adams, P.D., Keasling, J.D., Petzold, C.J., and Lee, T.S. (2013). Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production. Metab. Eng. *19*, 33–41.

Anderson, S.L., and McIntosh, L. (1991). Light-activated heterotrophic growth of the cyanobacterium *Synechocystis* sp. strain PCC 6803: a blue-light-requiring process. J. Bacteriol. *173*, 2761–2767.

Anfelt, J., Hallström, B., Nielsen, J., Uhlén, M., and Hudson, E.P. (2013). Using transcriptomics to improve butanol tolerance of *Synechocystis* sp. Strain PCC 6803. Appl. Environ. Microbiol. *79*, 7419–7427.

Anfelt, J., Kaczmarzyk, D., Shabestary, K., Renberg, B., Rockberg, J., Nielsen, J., Uhlén, M., and Hudson, E.P. (2015). Genetic and nutrient modulation of acetyl-CoA levels in *Synechocystis* for n-butanol production. Microb. Cell Factories *14*, 167.

Angermayr, S.A., Paszota, M., and Hellingwerf, K.J. (2012). Engineering a cyanobacterial cell factory for production of lactic acid. Appl. Environ. Microbiol. *78*, 7098–7106.

Atsumi, S., Higashide, W., and Liao, J.C. (2009). Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. Nat. Biotechnol. *27*, 1177–1180.

Barten, R., and Lill, H. (1995). DNA-uptake in the naturally competent cyanobacterium, *Synechocystis* sp. PCC 6803. FEMS Microbiol. Lett. *129*, 83–88.

Bauer, C.C., and Haselkorn, R. (1995). Vectors for determining the differential expression of genes in heterocysts and vegetative cells of *Anabaena* sp. strain PCC 7120. J. Bacteriol. *177*, 3332–3336.

Begemann, M.B., Zess, E.K., Walters, E.M., Schmitt, E.F., Markley, A.L., and Pfleger, B.F. (2013). An organic acid based counter selection system for cyanobacteria. PloS One *8*, e76594.

Behrendorff, J.B., Vickers, C.E., Chrysanthopoulos, P., and Nielsen, L.K. (2013). 2,2-Diphenyl-1-picrylhydrazyl as a screening tool for recombinant monoterpene biosynthesis. Microb. Cell Factories *12*, 76.

Bentley, F.K., García-Cerdán, J.G., Chen, H.-C., and Melis, A. (2013a). Paradigm of monoterpene (β-phellandrene) hydrocarbons production via photosynthesis in cyanobacteria. BioEnergy Res. *6*, 917–929.

Bentley, F.K., Zurbriggen, A., and Melis, A. (2013b). Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene. Mol. Plant.

Bishop, C.L. (2002). An investigation of the small polypeptides of photosystem II. Ph.D. University College London (University of London).

Bohlmann, J., Steele, C.L., and Croteau, R. (1997). Monoterpene synthases from grand fir (Abies grandis). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase. J. Biol. Chem. *272*, 21784–21792.

BP (2016). BP Statistical Review of World Energy 2016.

Branco dos Santos, F., Du, W., and Hellingwerf, K.J. (2014). *Synechocystis* : not just a plug-bug for CO_2, but a green *E. coli*. Front. Bioeng. Biotechnol. *2*.

Brennan, T.C.R., Turner, C.D., Krömer, J.O., and Nielsen, L.K. (2012). Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*. Biotechnol. Bioeng. *109*, 2513–2522.

Brennan, T.C.R., Kromer, J.O., and Nielsen, L.K. (2013). Physiological and transcriptional responses of *Saccharomyces cerevisiae* to d-limonene show changes to the cell wall but not to the plasma membrane. Appl. Environ. Microbiol. *79*, 3590–3600.

Bryant, D.A. (1982). Phycoerythrocyanin and phycoerythrin: properties and occurrence in cyanobacteria. Microbiology *128*, 835–844.

Byun-McKay, A., Godard, K.-A., Toudefallah, M., Martin, D.M., Alfaro, R., King, J., Bohlmann, J., and Plant, A.L. (2006). Wound-induced terpene synthase gene expression in Sitka spruce that exhibit resistance or susceptibility to attack by the white pine weevil. Plant Physiol. *140*, 1009–1021.

Cai, Y.P., and Wolk, C.P. (1990). Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. J. Bacteriol. *172*, 3138–3145.

Camsund, D., Heidorn, T., and Lindblad, P. (2014). Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. J. Biol. Eng. *8*, 4.

Capone, D.G. (2001). Marine nitrogen fixation: what's the fuss? Curr. Opin. Microbiol. *4*, 341–348.

Carter, O.A., Peters, R.J., and Croteau, R. (2003). Monoterpene biosynthesis pathway construction in *Escherichia coli*. Phytochemistry *64*, 425–433.

Čelešnik, H., Tanšek, A., Tahirović, A., Vižintin, A., Mustar, J., Vidmar, V., and Dolinar, M. (2016). Biosafety of biotechnologically important microalgae: intrinsic suicide switch implementation in cyanobacterium *Synechocystis* sp. PCC 6803. Biol. Open *5*, 519–528.

Cheah, Y.E., Albers, S.C., and Peebles, C.A.M. (2013). A novel counter-selection method for markerless genetic modification in *Synechocystis* sp. PCC 6803. Biotechnol. Prog. *29*, 23–30.

Chen, M., and Blankenship, R.E. (2011). Expanding the solar spectrum used by photosynthesis. Trends Plant Sci. *16*, 427–431.

Chen, G., Qu, S., Wang, Q., Bian, F., Peng, Z., Zhang, Y., Ge, H., Yu, J., Xuan, N., Bi, Y., et al. (2014). Transgenic expression of delta-6 and delta-15 fatty acid desaturases enhances omega-3 polyunsaturated fatty acid accumulation in *Synechocystis* sp. PCC6803. Biotechnol. Biofuels *7*, 32.

Chubukov, V., Mingardon, F., Schackwitz, W., Baidoo, E.E.K., Alonso-Gutierrez, J., Hu, Q., Lee, T.S., Keasling, J.D., and Mukhopadhyay, A. (2015). Acute limonene toxicity in *Escherichia coli* is caused by limonene-hydroperoxide and alleviated by a point mutation in alkyl hydroperoxidase (AhpC). Appl. Environ. Microbiol. AEM.01102-15.

Chuck, C.J., and Donnelly, J. (2014). The compatibility of potential bioderived fuels with Jet A-1 aviation kerosene. Appl. Energy *118*, 83–91.

Chungjatupornchai, W. (1990). Expression of the mosquitocidal-protein genes of *Bacillus thuringiensis* subsp. israelensis and the herbicide-resistance genebar in *Synechocystis* PCC6803. Curr. Microbiol. *21*, 283–288.

Colby, S.M., Alonso, W.R., Katahira, E.J., McGarvey, D.J., and Croteau, R. (1993). 4S-limonene synthase from the oil glands of spearmint (*Mentha spicata*). cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpene cyclase. J. Biol. Chem. *268*, 23016–23024.

Cot, S.S.-W., So, A.K.-C., and Espie, G.S. (2008). A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. J. Bacteriol. *190*, 936–945.

Cotton, C.A.R., Douglass, J.S., De Causmaecker, S., Brinkert, K., Cardona, T., Fantuzzi, A., Rutherford, A.W., and Murray, J.W. (2015). Photosynthetic constraints on fuel from microbes. Front. Bioeng. Biotechnol. *3*.

Danielsen, S., Kilstrup, M., Barilla, K., Jochimsen, B., and Neuhard, J. (1992). Characterization of the *Escherichia coli codBA* operon encoding cytosine permease and cytosine deaminase. Mol. Microbiol. *6*, 1335–1344.

Davies, F.K., Work, V.H., Beliaev, A.S., and Posewitz, M.C. (2014). engineering limonene and bisabolene production in wild type and a glycogen-deficient mutant of *Synechococcus* sp. PCC 7002. Front. Bioeng. Biotechnol. *2*.

Debuchy, R., Purton, S., and Rochaix, J.D. (1989). The argininosuccinate lyase gene of C*hlamydomonas reinhardtii:* an important tool for nuclear transformation and for correlating the genetic and molecular maps of the ARG7 locus. EMBO J. *8*, 2803–2809.

Deng, M.-D., and Coleman, J.R. (1999). Ethanol synthesis by genetic engineering in cyanobacteria. Appl. Environ. Microbiol. *65*, 523–528.

Dexter, J., and Fu, P. (2009). Metabolic engineering of cyanobacteria for ethanol production. Energy Environ. Sci. *2*, 857–864.

Du, W., Liang, F., Duan, Y., Tan, X., and Lu, X. (2013). Exploring the photosynthetic production capacity of sucrose by cyanobacteria. Metab. Eng. *19*, 17–25.

Dubeau, M.-P., Ghinet, M.G., Jacques, P.-E., Clermont, N., Beaulieu, C., and Brzezinski, R. (2009). Cytosine deaminase as a negative selection marker for gene disruption and replacement in the genus streptomyces and other actinobacteria. Appl. Environ. Microbiol. *75*, 1211–1214.

Ducat, D.C., Way, J.C., and Silver, P.A. (2011a). Engineering cyanobacteria to generate high-value products. Trends Biotechnol. *29*, 95–103.

Ducat, D.C., Sachdeva, G., and Silver, P.A. (2011b). Rewiring hydrogenase-dependent redox circuits in cyanobacteria. Proc. Natl. Acad. Sci. *108*, 3941–3946.

Ducat, D.C., Avelar-Rivas, J.A., Way, J.C., and Silver, P.A. (2012). Rerouting carbon flux to enhance photosynthetic productivity. Appl. Environ. Microbiol. *78*, 2660–2668.

Dühring, U., Axmann, I.M., Hess, W.R., and Wilde, A. (2006). An internal antisense RNA regulates expression of the photosynthesis gene isiA. Proc. Natl. Acad. Sci. U. S. A. *103*, 7054–7058.

Dunlop, M.J. (2011). Engineering microbes for tolerance to next-generation biofuels. Biotechnol. Biofuels *4*, 32.

Dunlop, M.J., Dossani, Z.Y., Szmidt, H.L., Chu, H.C., Lee, T.S., Keasling, J.D., Hadi, M.Z., and Mukhopadhyay, A. (2011). Engineering microbial biofuel tolerance and export using efflux pumps. Mol. Syst. Biol. *7*, 487.

Durall, C., and Lindblad, P. (2015). Mechanisms of carbon fixation and engineering for increased carbon fixation in cyanobacteria. Algal Res. *11*, 263–270.

Durão, P., Aigner, H., Nagy, P., Mueller-Cajar, O., Hartl, F.U., and Hayer-Hartl, M. (2015). Opposing effects of folding and assembly chaperones on evolvability of rubisco. Nat. Chem. Biol. *11*, 148–155.

Elhai, J. (1994). Genetic techniques appropriate for the biotechnological exploitation of cyanobacteria. J. Appl. Phycol. *6*, 177–186.

Elhai, J., and Wolk, C.P. (1988). Conjugal transfer of DNA to cyanobacteria. Methods Enzymol. *167*, 747–754.

Englund, E., Pattanaik, B., Ubhayasekera, S.J.K., Stensjö, K., Bergquist, J., and Lindberg, P. (2014). Production of squalene in *Synechocystis* sp. PCC 6803. PLoS ONE *9*, e90270.

Entova, S. (2013). Kinetic characterization, crystallization, and photosynthetic expression of (+)-4R-limonene synthase from C. sinensis. MS. Brandeis University.

Erdrich, P., Knoop, H., Steuer, R., and Klamt, S. (2014). Cyanobacterial biofuels: new insights and strain design strategies revealed by computational modeling. Microb. Cell Factories *13*, 128.

Erickson, B., Nelson, J.E., and Winters, P. (2012). Perspective on opportunities in industrial biotechnology in renewable chemicals. Biotechnol. J. *7*, 176–185.

Flores, E., and Wolk, C.P. (1985). Identification of facultatively heterotrophic, N_2-fixing cyanobacteria able to receive plasmid vectors from *Escherichia coli* by conjugation. J. Bacteriol. *162*, 1339–1341.

Fu, P. (2009). Genome-scale modeling of *Synechocystis* sp. PCC 6803 and prediction of pathway insertion. J. Chem. Technol. Biotechnol. *84*, 473–483.

Fuchita, M., Ardiani, A., Zhao, L., Serve, K., Stoddard, B.L., and Black, M.E. (2009). Bacterial cytosine deaminase mutants created by molecular engineering demonstrate improved 5fc-mediated cell killing in vitro and in vivo. Cancer Res. *69*, 4791–4799.

García-Domínguez, M., Lopez-Maury, L., Florencio, F.J., and Reyes, J.C. (2000). A gene cluster involved in metal homeostasis in the cyanobacterium *Synechocystis* sp. strain PCC 6803. J. Bacteriol. *182*, 1507–1514.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. Nature *403*, 339–342.

Gay, P., Le Coq, D., Steinmetz, M., Ferrari, E., and Hoch, J.A. (1983). Cloning structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. J. Bacteriol. *153*, 1424–1431.

Geerts, D., Bovy, A., de Vrieze, G., Borrias, M., and Weisbeek, P. (1995). Inducible expression of heterologous genes targeted to a chromosomal platform in the cyanobacterium *Synechococcus* sp. PCC 7942. Microbiology *141*, 831–841.

Gill, R.T., Katsoulakis, E., Schmitt, W., Taroncher-Oldenburg, G., Misra, J., and Stephanopoulos, G. (2002). Genome-wide dynamic transcriptional profiling of the light-to-dark transition in *Synechocystis* sp. Strain PCC 6803. J. Bacteriol. *184*, 3671–3681.

Gillor, O., Harush, A., Hadas, O., Post, A.F., and Belkin, S. (2003). A *Synechococcus* PglnA::*luxAB* fusion for estimation of nitrogen bioavailability to freshwater cyanobacteria. Appl. Environ. Microbiol. *69*, 1465–1474.

Gordon, G.C., Korosh, T.C., Cameron, J.C., Markley, A.L., Begemann, M.B., and Pfleger, B.F. (2016). CRISPR interference as a titratable, trans-acting regulatory tool for metabolic engineering in the cyanobacterium *Synechococcus* sp. strain PCC 7002. Metab. Eng. *38*, 170–179.

Griese, M., Lange, C., and Soppa, J. (2011). Ploidy in cyanobacteria. FEMS Microbiol. Lett. *323*, 124–131.

Grigorieva, G., and Shestakov, S. (1982). Transformation in the cyanobacterium *Synechocystis* sp. 6803. FEMS Microbiol. Lett. *13*, 367–370.

Grossman, A.R., Schaefer, M.R., Chiang, G.G., and Collier, J.L. (1993). The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiol. Rev. *57*, 725–749.

Guerrero, F., Carbonell, V., Cossu, M., Correddu, D., and Jones, P.R. (2012). Ethylene synthesis and regulated expression of recombinant protein in *Synechocystis* sp. PCC 6803. PLOS ONE *7*, e50470.

Halfmann, C., Gu, L., Gibbons, W., and Zhou, R. (2014a). Genetically engineering cyanobacteria to convert CO2, water, and light into the long-chain hydrocarbon farnesene. Appl. Microbiol. Biotechnol. *98*, 9869–9877.

Halfmann, C., Gu, L., and Zhou, R. (2014b). Engineering cyanobacteria for the production of a cyclic hydrocarbon fuel from CO_2 and H_2O. Green Chem. *16*, 3175–3185.

Han, G.H., Gim, G.H., Kim, W., Seo, S.I., and Kim, S.W. (2013). Enhanced indirubin production in recombinant *Escherichia coli* harboring a flavin-containing monooxygenase gene by cysteine supplementation. J. Biotechnol. *164*, 179–187.

He, Q., Schlich, T., Paulsen, H., and Vermaas, W. (1999). Expression of a higher plant light-harvesting chlorophyll a/b-binding protein in *Synechocystis* sp. PCC 6803. Eur. J. Biochem. FEBS *263*, 561–570.

Heidorn, T., Camsund, D., Huang, H.-H., Lindberg, P., Oliveira, P., Stensjö, K., and Lindblad, P. (2011). Chapter Twenty-Four - Synthetic biology in cyanobacteria: engineering and analyzing novel functions. In Methods in Enzymology, C. Voigt, ed. (Academic Press), pp. 539–579.

Heipieper, H.J., Weber, F.J., Sikkema, J., Keweloh, H., and de Bont, J.A.M. (1994). Mechanisms of resistance of whole cells to toxic organic solvents. Trends Biotechnol. *12*, 409–415.

Hernández, J.A., Muro-Pastor, A.M., Flores, E., Bes, M.T., Peleato, M.L., and Fillat, M.F. (2006). Identification of a *furA cis* antisense RNA in the cyanobacterium *Anabaena* sp. PCC 7120. J. Mol. Biol. *355*, 325–334.

Higo, A., Isu, A., Fukaya, Y., and Hisabori, T. (2016). Efficient gene induction and endogenous gene repression systems for the filamentous cyanobacterium *Anabaena* sp. PCC 7120. Plant Cell Physiol. *57*, 387–396.

Hihara, Y., and Ikeuchi, M. (1997). Mutation in a novel gene required for photomixotrophic growth leads to enhanced photoautotrophic growth of *Synechocystis* sp. PCC 6803. Photosynth. Res. *53*, 243–252.

Hirokawa, Y., Maki, Y., Tatsuke, T., and Hanai, T. (2016). Cyanobacterial production of 1,3-propanediol directly from carbon dioxide using a synthetic metabolic pathway. Metab. Eng. *34*, 97–103.

Holtman, C.K., Chen, Y., Sandoval, P., Gonzales, A., Nalty, M.S., Thomas, T.L., Youderian, P., and Golden, S.S. (2005). High-throughput functional analysis of the *Synechococcus elongatus* PCC 7942 genome. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes *12*, 103–115.

Huang, H.-H., and Lindblad, P. (2013). Wide-dynamic-range promoters engineered for cyanobacteria. J. Biol. Eng. *7*, 10.

Huang, H.-H., Camsund, D., Lindblad, P., and Heidorn, T. (2010). Design and characterization of molecular tools for a synthetic biology approach towards developing cyanobacterial biotechnology. Nucleic Acids Res. *38*, 2577–2593.

Ikeuchi, M., and Tabata, S. (2001). *Synechocystis* sp. PCC 6803 — a useful tool in the study of the genetics of cyanobacteria. Photosynth. Res. *70*, 73–83.

Imamura, S., and Asayama, M. (2009). Sigma factors for cyanobacterial transcription. Gene Regul. Syst. Biol. *3*, 65–87.

Immethun, C.M., Ng, K.M., DeLorenzo, D.M., Waldron-Feinstein, B., Lee, Y.-C., and Moon, T.S. (2016). Oxygen-responsive genetic circuits constructed in *Synechocystis* sp. PCC 6803. Biotechnol. Bioeng. *113*, 433–442.

International Energy Agency (2014). World Energy Outlook 2014 Executive Summary.

International Energy Agency (2015). Key World Energy Statistics 2015.

International Energy Agency (2016a). Key World Energy Trends Excerpt from: World Energy Balances.

International Energy Agency (2016b). Key Oil Trends Excerpt from: Oil Information.

Iwai, M., Katoh, H., Katayama, M., and Ikeuchi, M. (2004). Improved genetic transformation of the thermophilic cyanobacterium, *Thermosynechococcus elongatus* BP-1. Plant Cell Physiol. *45*, 171–175.

Jacobsen, J.H., and Frigaard, N.-U. (2014). Engineering of photosynthetic mannitol biosynthesis from CO_2 in a cyanobacterium. Metab. Eng. *21*, 60–70.

Jensen, R.G. (2000). Activation of rubisco regulates photosynthesis at high temperature and CO_2. Proc. Natl. Acad. Sci. U. S. A. *97*, 12937–12938.

Jensen, P.E., and Leister, D. (2014). Cyanobacteria as an experimental platform for modifying bacterial and plant photosynthesis. Front. Bioeng. Biotechnol. *2*.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Jongedijk, E., Cankar, K., Ranzijn, J., van der Krol, S., Bouwmeester, H., and Beekwilder, J. (2015). Capturing of the monoterpene olefin limonene produced in *Saccharomyces cerevisiae*. Yeast *32*, 159–171.

Jongedijk, E., Cankar, K., Buchhaupt, M., Schrader, J., Bouwmeester, H., and Beekwilder, J. (2016). Biotechnological production of limonene in microorganisms. Appl. Microbiol. Biotechnol. *100*, 2927–2938.

Kaczmarzyk, D., Anfelt, J., Särnegrim, A., and Hudson, E.P. (2014). Overexpression of sigma factor SigB improves temperature and butanol tolerance of *Synechocystis* sp. PCC6803. J. Biotechnol. *182–183*, 54–60.

Kaiser, B.K., Carleton, M., Hickman, J.W., Miller, C., Lawson, D., Budde, M., Warrener, P., Paredes, A., Mullapudi, S., Navarro, P., et al. (2013). Fatty aldehydes in cyanobacteria are a metabolically flexible precursor for a diversity of biofuel products. PLOS ONE *8*, e58307.

Kamennaya, N.A., Ahn, S., Park, H., Bartal, R., Sasaki, K.A., Holman, H.-Y., and Jansson, C. (2015). Installing extra bicarbonate transporters in the cyanobacterium *Synechocystis* sp. PCC6803 enhances biomass production. Metab. Eng. *29*, 76–85.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., et al. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. Strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. *3*, 109–136.

Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S., Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T., et al. (2001). Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes *8*, 205-213-253.

Kerfeld, C.A., and Melnicki, M.R. (2016). Assembly, function and evolution of cyanobacterial carboxysomes. Curr. Opin. Plant Biol. *31*, 66–75.

Kiyota, H., Okuda, Y., Ito, M., Hirai, M.Y., and Ikeuchi, M. (2014). Engineering of cyanobacteria for the photosynthetic production of limonene from CO2. J. Biotechnol. *185*, 1–7.

Kleiner, D. (1978). Inhibition of the respiratory system in *Azotobacter vinelandii* by divalent transition metal ions. FEBS Lett. *96*, 364–366.

Knoop, H., and Steuer, R. (2015). A computational analysis of stoichiometric constraints and trade-offs in cyanobacterial biofuel production. Front. Bioeng. Biotechnol. *3*.

Knoshaug, E.P., and Zhang, M. (2008). Butanol tolerance in a selection of microorganisms. Appl. Biochem. Biotechnol. *153*, 13–20.

Kobayashi, T., Hisajima, S., Stougaard, J., and Ichikawa, H. (1995). A conditional negative selection for *Arabidopsis* expressing a bacterial cytosine deaminase gene. Jpn. J. Genet. *70*, 409–422.

Koksharova, O.A., and Wolk, C.P. (2002). Genetic tools for cyanobacteria. Appl. Microbiol. Biotechnol. *58*, 123–137.

Kopf, M., and Hess, W.R. (2015). Regulatory RNAs in photosynthetic cyanobacteria. FEMS Microbiol. Rev. *39*, 301–315.

Kufryk, G.I., Sachet, M., Schmetterer, G., and Vermaas, W.F.J. (2002). Transformation of the cyanobacterium *Synechocystis* sp. PCC 6803 as a tool for genetic mapping: optimization of efficiency. FEMS Microbiol. Lett. *206*, 215–219.

Kuhlemeier, C.J., Thomas, A.A.M., van der Ende, A., van Leen, R.W., Borrias, W.E., van den Hondel, C.A.M.J.J., and van Arkel, G.A. (1983). A host-vector system for gene cloning in the cyanobacterium *Anacystis nidulans* R2. Plasmid *10*, 156–163.

Kunert, A., Hagemann, M., and Erdmann, N. (2000). Construction of promoter probe vectors for *Synechocystis* sp. PCC 6803 using the light-emitting reporter systems Gfp and LuxAB. J. Microbiol. Methods *41*, 185–194.

Kusakabe, T., Tatsuke, T., Tsuruno, K., Hirokawa, Y., Atsumi, S., Liao, J.C., and Hanai, T. (2013). Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. Metab. Eng. *20*, 101–108.

Labarre, J., Chauvat, F., and Thuriaux, P. (1989). Insertional mutagenesis by random cloning of antibiotic resistance genes into the genome of the cyanobacterium *Synechocystis* strain PCC 6803. J. Bacteriol. *171*, 3449–3457.

Lagarde, D., Beuf, L., and Vermaas, W. (2000). Increased production of zeaxanthin and other pigments by application of genetic engineering techniques to *Synechocystis* sp. Strain PCC 6803. Appl. Environ. Microbiol. *66*, 64–72.

Lan, E.I., and Liao, J.C. (2011). Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. Metab. Eng. *13*, 353–363.

Landry, B.P., Stöckel, J., and Pakrasi, H.B. (2013). Use of degradation tags to control protein levels in the cyanobacterium *Synechocystis* sp. strain PCC 6803. Appl. Environ. Microbiol. *79*, 2833–2835.

Lange, B.M. (2015). Biosynthesis and biotechnology of high-value p-menthane monoterpenes, including menthol, carvone, and limonene. In Biotechnology of Isoprenoids, J. Schrader, and J. Bohlmann, eds. (Springer International Publishing), pp. 319–353.

Lea-Smith, D.J., Ross, N., Zori, M., Bendall, D.S., Dennis, J.S., Scott, S.A., Smith, A.G., and Howe, C.J. (2013). Thylakoid terminal oxidases are essential for the cyanobacterium *Synechocystis* sp. PCC 6803 to survive rapidly changing light intensities1[C][W][OA]. Plant Physiol. *162*, 484–495.

Lea-Smith, D.J., Bombelli, P., Dennis, J.S., Scott, S.A., Smith, A.G., and Howe, C.J. (2014). Phycobilisome-deficient strains of *Synechocystis* sp. PCC 6803 have reduced size and require carbon-limiting conditions to exhibit enhanced productivity [W][OPEN]. Plant Physiol. *165*, 705–714.

Lee, S.K., Chou, H., Ham, T.S., Lee, T.S., and Keasling, J.D. (2008). Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. Curr. Opin. Biotechnol. *19*, 556–563.

Li, H., and Liao, J.C. (2013). Engineering a cyanobacterium as the catalyst for the photosynthetic conversion of CO_2 to 1,2-propanediol. Microb. Cell Factories *12*, 4.

Liang, F., and Lindblad, P. (2016). Effects of overexpressing photosynthetic carbon flux control enzymes in the cyanobacterium *Synechocystis* PCC 6803. Metab. Eng. *38*, 56–64.

Liao, J.C., Mi, L., Pontrelli, S., and Luo, S. (2016). Fuelling the future: microbial engineering for the production of sustainable biofuels. Nat. Rev. Microbiol. *14*, 288–304.

Liberton, M., Berg, R.H., Heuser, J., Roth, R., and Pakrasi, H.B. (2006). Ultrastructure of the membrane systems in the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. Protoplasma *227*, 129.

Lindberg, P., Park, S., and Melis, A. (2010). Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. Metab. Eng. *12*, 70–79.

Lindblad, P., Lindberg, P., Oliveira, P., Stensjö, K., and Heidorn, T. (2012). Design, engineering, and construction of photosynthetic microbial cell factories for renewable solar fuel production. Ambio *41*, 163–168.

Liu, X., and Curtiss, R. (2009). Nickel-inducible lysis system in *Synechocystis* sp. PCC 6803. Proc. Natl. Acad. Sci. *106*, 21550–21554.

Liu, J., Chen, L., Wang, J., Qiao, J., and Zhang, W. (2012). Proteomic analysis reveals resistance mechanism against biofuel hexane in *Synechocystis* sp. PCC 6803. Biotechnol. Biofuels *5*, 68.

Liu, X., Sheng, J., and Curtiss III, R. (2011a). Fatty acid production in genetically modified cyanobacteria. Proc. Natl. Acad. Sci. U. S. A. *108*, 6899–6904.

Liu, X., Fallon, S., Sheng, J., and Curtiss, R. (2011b). CO_2-limitation-inducible green recovery of fatty acids from cyanobacterial biomass. Proc. Natl. Acad. Sci. *108*, 6905–6908.

Lluisma, A., Karmacharya, N., Zarka, A., Ben-Dov, E., Zaritsky, A., and Boussiba, S. (2001). Suitability of *Anabaena* PCC7120 expressing mosquitocidal toxin genes from *Bacillus thuringiensis* subsp. israelensis for biotechnological application. Appl. Microbiol. Biotechnol. *57*, 161–166.

López-Maury, L., García-Domínguez, M., Florencio, F.J., and Reyes, J.C. (2002). A two-component signal transduction system involved in nickel sensing in the cyanobacterium *Synechocystis* sp. PCC 6803. Mol. Microbiol. *43*, 247–256.

Lücker, J., El Tamer, M.K., Schwab, W., Verstappen, F.W.A., van der Plas, L.H.W., Bouwmeester, H.J., and Verhoeven, H.A. (2002). Monoterpene biosynthesis in lemon (*Citrus limon*). cDNA isolation and functional analysis of four monoterpene synthases. Eur. J. Biochem. FEBS *269*, 3160–3171.

Ma, A.T., Schmidt, C.M., and Golden, J.W. (2014). Regulation of gene expression in diverse cyanobacterial species by using theophylline-responsive riboswitches. Appl. Environ. Microbiol. *80*, 6704–6713.

Ma, J., Campbell, A., and Karlin, S. (2002). Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. J. Bacteriol. *184*, 5733–5745.

Machado, I.M.P., and Atsumi, S. (2012). Cyanobacterial biofuel production. J. Biotechnol. *162*, 50–56.

Mackrow, B. (2011). Employing a synthetic biology "parts-based" approach to algal biotechnology. MRes. University College London (University of London).

Markley, A.L., Begemann, M.B., Clarke, R.E., Gordon, G.C., and Pfleger, B.F. (2015). Synthetic biology toolbox for controlling gene expression in the cyanobacterium *Synechococcus* sp. strain PCC 7002. ACS Synth. Biol. *4*, 595–603.

Marraccini, P., Bulteau, S., Cassier-Chauvat, C., Mermet-Bouvier, P., and Chauvat, F. (1993). A conjugative plasmid vector for promoter analysis in several cyanobacteria of the genera *Synechococcus* and *Synechocystis* . Plant Mol. Biol. *23*, 905–909.

Maruyama, T., Ito, M., Kiuchi, F., and Honda, G. (2001). Molecular cloning, functional expression and characterization of d-limonene synthase from *Schizonepeta tenuifolia*. Biol. Pharm. Bull. *24*, 373–377.

Maruyama, T., Saeki, D., Ito, M., and Honda, G. (2002). Molecular cloning, functional expression and characterization of d-limonene synthase from Agastache rugosa. Biol. Pharm. Bull. *25*, 661–665.

Matsunaga, T., and Takeyama, H. (1994). Genetic engineering in marine cyanobacteria. J. Appl. Phycol. *7*, 77–84.

Matsuoka, M., Takahama, K., and Ogawa, T. (2001). Gene replacement in cyanobacteria mediated by a dominant streptomycin-sensitive rps12 gene that allows selection of mutants free from drug resistance markers. Microbiol. Read. Engl. *147*, 2077–2087.

McFadden, G.I. (2001). Chloroplast origin and integration. Plant Physiol. *125*, 50–53.

McKenna, R., and Nielsen, D.R. (2011). Styrene biosynthesis from glucose by engineered E. coli. Metab. Eng. *13*, 544–554.

McKenna, R., Pugh, S., Thompson, B., and Nielsen, D.R. (2013). Microbial production of the aromatic building-blocks (S)-styrene oxide and (R)-1,2-phenylethanediol from renewable resources. Biotechnol. J. *8*, 1465–1475.

McKenna, R., Moya, L., McDaniel, M., and Nielsen, D.R. (2014). Comparing in situ removal strategies for improving styrene bioproduction. Bioprocess Biosyst. Eng. *38*, 165–174.

Meyer, A., Würsten, M., Schmid, A., Kohler, H.-P.E., and Witholt, B. (2002). Hydroxylation of indole by laboratory-evolved 2-hydroxybiphenyl 3-monooxygenase. J. Biol. Chem. *277*, 34161–34167.

Mitschke, J., Georg, J., Scholz, I., Sharma, C.M., Dienst, D., Bantscheff, J., Voss, B., Steglich, C., Wilde, A., Vogel, J., et al. (2011). An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. Proc. Natl. Acad. Sci. U. S. A. *108*, 2124–2129.

Miyake, M., Takase, K., Narato, M., Khatipov, E., Schnackenberg, J., Shirai, M., Kurane, R., and Asada, Y. (2000). Polyhydroxybutyrate production from carbon dioxide by cyanobacteria. Appl. Biochem. Biotechnol. *84–86*, 991.

Mohamed, A., Eriksson, J., Osiewacz, H.D., and Jansson, C. (1993). Differential expression of the *psbA* genes in the cyanobacterium *Synechocystis* 6803. Mol. Gen. Genet. MGG *238*, 161–168.

Montagud, A., Navarro, E., Fernández de Córdoba, P., Urchueguía, J.F., and Patil, K.R. (2010). Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. BMC Syst. Biol. *4*, 156.

Montagud, A., Zelezniak, A., Navarro, E., de Córdoba, P.F., Urchueguía, J.F., and Patil, K.R. (2011). Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium *Synechocystis* sp. PCC6803. Biotechnol. J. *6*, 330–342.

Mooney, A., Ward, P.G., and O'Connor, K.E. (2006). Microbial degradation of styrene: biochemistry, molecular genetics, and perspectives for biotechnological applications. Appl. Microbiol. Biotechnol. *72*, 1–10.

Moreno-Hagelsieb, G., and Collado-Vides, J. (2002). A powerful non-homology method for the prediction of operons in prokaryotes. Bioinformatics *18*, S329–S336.

Moser, D., Zarka, D., Hedman, C., and Kallas, T. (1995). Plasmid and chromosomal DNA recovery by electroextraction of cyanobacteria. FEMS Microbiol. Lett. *128*, 307–313.

Moser, F., Broers, N.J., Hartmans, S., Tamsir, A., Kerkman, R., Roubos, J.A., Bovenberg, R., and Voigt, C.A. (2012). Genetic circuit performance under conditions relevant for industrial bioreactors. ACS Synth. Biol. *1*, 555–564.

Mubarak, M., Shaija, A., and Suchithra, T.V. (2015). A review on the extraction of lipid from microalgae for biodiesel production. Algal Res. *7*, 117–123.

Mueller, T.J., Berla, B.M., Pakrasi, H.B., and Maranas, C.D. (2013). Rapid construction of metabolic models for a family of Cyanobacteria using a multiple source annotation workflow. BMC Syst. Biol. *7*, 142.

Mühlenhoff, U., and Chauvat, F. (1996). Gene transfer and manipulation in the thermophilic cyanobacterium*Synechococcus elongatus*. Mol. Gen. Genet. MGG *252*, 93–100.

Muramatsu, M., and Hihara, Y. (2006). Characterization of high-light-responsive promoters of the *psaAB* genes in *Synechocystis* sp. PCC 6803. Plant Cell Physiol. *47*, 878–890.

Muramatsu, M., and Hihara, Y. (2007). Coordinated high-light response of genes encoding subunits of photosystem IIs achieved by AT-rich upstream sequences in the cyanobacterium *Synechocystis* sp. strain PCC 6803. J. Bacteriol. *189*, 2750–2758.

Nakahira, Y., Ogawa, A., Asano, H., Oyama, T., and Tozawa, Y. (2013). Theophylline-dependent riboswitch as a novel genetic tool for strict regulation of protein expression in cyanobacterium *Synechococcus elongatus* PCC 7942. Plant Cell Physiol. *54*, 1724–1735.

Nakamura, Y., Kaneko, T., Hirosawa, M., Miyajima, N., and Tabata, S. (1998). CyanoBase, a www database containing the complete nucleotide sequence of the genome of *Synechocystis* sp. strain PCC6803. Nucleic Acids Res. *26*, 63–67.

Nakamura, Y., Kaneko, T., Miyajima, N., and Tabata, S. (1999). Extension of CyanoBase. CyanoMutants: repository of mutant information on *Synechocystis* sp. strain PCC6803. Nucleic Acids Res. *27*, 66–68.

Nakamura, Y., Kaneko, T., Sato, S., Ikeuchi, M., Katoh, H., Sasamoto, S., Watanabe, A., Iriguchi, M., Kawashima, K., Kimura, T., et al. (2002). Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes *9*, 123–130.

Newman, J.D., Marshall, J., Chang, M., Nowroozi, F., Paradise, E., Pitera, D., Newman, K.L., and Keasling, J.D. (2006). High-level production of amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered *Escherichia coli*. Biotechnol. Bioeng. *95*, 684–691.

Ng, A.H., Berla, B.M., and Pakrasi, H.B. (2015). Fine-tuning of photoautotrophic protein production by combining promoters and neutral sites in the cyanobacterium *Synechocystis* sp. Strain PCC 6803. Appl. Environ. Microbiol. *81*, 6857–6863.

Ng, W.O., Zentella, R., Wang, Y., Taylor, J.S., and Pakrasi, H.B. (2000). PhrA, the major photoreactivating factor in the cyanobacterium *Synechocystis* sp. strain PCC 6803 codes for a cyclobutane-pyrimidine-dimer-specific DNA photolyase. Arch. Microbiol. *173*, 412–417.

Nickelsen, J., Rengstl, B., Stengel, A., Schottkowski, M., Soll, J., and Ankele, E. (2011). Biogenesis of the cyanobacterial thylakoid membrane system – an update. FEMS Microbiol. Lett. *315*, 1–5.

Niederholtmeyer, H., Wolfstädter, B.T., Savage, D.F., Silver, P.A., and Way, J.C. (2010). Engineering cyanobacteria to synthesize and export hydrophilic products. Appl. Environ. Microbiol. *76*, 3462–3466.

Nishio, T., Patel, A., Wang, Y., and Lau, P.C. (2001). Biotransformations catalyzed by cloned p-cymene monooxygenase from Pseudomonas putida F1. Appl. Microbiol. Biotechnol. *55*, 321–325.

Nishioka, M., Nakai, K., Miyake, M., Asada, Y., and Taya, M. (2001). Production of poly-β-hydroxybutyrate by thermophilic cyanobacterium, *Synechococcus* sp. MA19, under phosphate-limited conditions. Biotechnol. Lett. *23*, 1095–1099.

Niu, X., Zhu, Y., Pei, G., Wu, L., Chen, L., and Zhang, W. (2015). Elucidating butanol tolerance mediated by a response regulator Sll0039 in *Synechocystis* sp. PCC 6803 using a metabolomic approach. Appl. Microbiol. Biotechnol. *99*, 1845–1857.

Nobles, D.R., and Brown, R.M. (2008). Transgenic expression of *Gluconacetobacter xylinus* strain ATCC 53582 cellulose synthase genes in the cyanobacterium *Synechococcus leopoliensis* strain UTCC 100. Cellulose *15*, 691–701.

Nogales, J., Gudmundsson, S., and Thiele, I. (2013). Toward systems metabolic engineering in cyanobacteria: opportunities and bottlenecks. Bioengineered *4*, 158–163.

Nöthe, C., and Hartmans, S. (1994). Formation and degradation of styrene oxide stereoisomers by different microorganisms. Biocatalysis *10*, 219–225.

Nozzi, N.E., Oliver, J.W.K., and Atsumi, S. (2013). Cyanobacteria as a platform for biofuel production. Synth. Biol. *1*, 7.

Ohbayashi, R., Akai, H., Yoshikawa, H., Hess, W.R., and Watanabe, S. (2016). A tightly inducible riboswitch system in *Synechocystis* sp. PCC 6803. J. Gen. Appl. Microbiol. *62*, 154–159.

Okada, K., Minehira, M., Zhu, X., Suzuki, K., Nakagawa, T., Matsuda, H., and Kawamukai, M. (1997). The *ispB* gene encoding octaprenyl diphosphate synthase is essential for growth of *Escherichia coli*. J. Bacteriol. *179*, 3058–3060.

Oliver, J.W.K., Machado, I.M.P., Yoneda, H., and Atsumi, S. (2013). Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. Proc. Natl. Acad. Sci. U. S. A. *110*, 1249–1254.

Onai, K., Morishita, M., Kaneko, T., Tabata, S., and Ishiura, M. (2003). Natural transformation of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1: a simple and efficient method for gene transfer. Mol. Genet. Genomics *271*, 50–59.

Otto, K., Hofstetter, K., Röthlisberger, M., Witholt, B., and Schmid, A. (2004). Biochemical characterization of StyAB from *Pseudomonas* sp. strain VLB120 as a two-component flavin-diffusible monooxygenase. J. Bacteriol. *186*, 5292–5302.

Panda, B., and Mallick, N. (2007). Enhanced poly-β-hydroxybutyrate accumulation in a unicellular cyanobacterium, *Synechocystis* sp. PCC 6803. Lett. Appl. Microbiol. *44*, 194–198.

Panke, S., Witholt, B., Schmid, A., and Wubbolts, M.G. (1998). Towards a biocatalyst for (S)-styrene oxide production: characterization of the styrene degradation pathway of *Pseudomonas* sp. strain VLB120. Appl. Environ. Microbiol. *64*, 2032–2043.

Panke, S., Wubbolts, M.G., Schmid, A., and Witholt, B. (2000). Production of enantiopure styrene oxide by recombinant *Escherichia coli* synthesizing a two-component styrene monooxygenase. Biotechnol. Bioeng. *69*, 91–100.

Panke, S., Held, M., Wubbolts, M.G., Witholt, B., and Schmid, A. (2002). Pilot-scale production of (S)-styrene oxide from styrene by recombinant *Escherichia coli* synthesizing styrene monooxygenase. Biotechnol. Bioeng. *80*, 33–41.

Parmar, A., Singh, N.K., Pandey, A., Gnansounou, E., and Madamwar, D. (2011). Cyanobacteria and microalgae: A positive prospect for biofuels. Bioresour. Technol. *102*, 10163–10172.

Pattanaik, B., and Lindberg, P. (2015). Terpenoids and their biosynthesis in cyanobacteria. Life *5*, 269–293.

Peca, L., Kós, P.B., Máté, Z., Farsang, A., and Vass, I. (2008). Construction of bioluminescent cyanobacterial reporter strains for detection of nickel, cobalt and zinc. FEMS Microbiol. Lett. *289*, 258–264.

Peralta-Yahya, P.P., Zhang, F., del Cardayre, S.B., and Keasling, J.D. (2012). Microbial engineering for the production of advanced biofuels. Nature *488*, 320–328.

Porta, D., Bullerjahn, G.S., Durham, K.A., Wilhelm, S.W., Twiss, M.R., and McKay, R.M.L. (2003). Physiological characterization of a *Synechococcus* Sp. (cyanophyceae) strain PCC 7942 iron-dependent bioreporter for freshwater environments1. J. Phycol. *39*, 64–73.

Price, M.N., Huang, K.H., Alm, E.J., and Arkin, A.P. (2005). A novel method for accurate operon predictions in all sequenced prokaryotes. Nucleic Acids Res. *33*, 880–892.

Qiao, J., Wang, J., Chen, L., Tian, X., Huang, S., Ren, X., and Zhang, W. (2012). Quantitative iTRAQ LC-MS/MS proteomics reveals metabolic responses to biofuel ethanol in cyanobacterial *Synechocystis* sp. PCC 6803. J. Proteome Res. *11*, 5286–5300.

Quax, T.E.F., Claassens, N.J., Söll, D., and van der Oost, J. (2015). Codon Bias as a Means to Fine-Tune Gene Expression. Mol. Cell *59*, 149–161.

Quintana, N., Van der Kooy, F., Van de Rhee, M.D., Voshol, G.P., and Verpoorte, R. (2011). Renewable energy from Cyanobacteria: energy production optimization by metabolic pathway engineering. Appl. Microbiol. Biotechnol. *91*, 471–490.

Rai, P.K., Mallick, N., and Rai, L.C. (1994). Effect of Cu and Ni on growth, mineral uptake, photosynthesis and enzyme activities of *Chlorella vulgaris* at different pH values. Biomed. Environ. Sci. BES *7*, 56–67.

Ramos, J.L., Duque, E., Gallegos, M.-T., Godoy, P., Ramos-Gonzalez, M.I., Rojas, A., Teran, W., and Segura, A. (2002). Mechanisms of solvent tolerance in gram-negative bacteria. Annu. Rev. Microbiol. *56*, 743–768.

Rasmussen, B., Fletcher, I.R., Brocks, J.J., and Kilburn, M.R. (2008). Reassessing the first appearance of eukaryotes and cyanobacteria. Nature *455*, 1101–1104.

Reinsvold, R.E., Jinkerson, R.E., Radakovits, R., Posewitz, M.C., and Basu, C. (2011). The production of the sesquiterpene β-caryophyllene in a transgenic strain of the cyanobacterium *Synechocystis* . J. Plant Physiol. *168*, 848–852.

Rice, D., Mazur, B.J., and Haselkorn, R. (1982). Isolation and physical mapping of nitrogen fixation genes from the cyanobacterium *Anabaena* 7120. J. Biol. Chem. *257*, 13157–13163.

Ried, J.L., and Collmer, A. (1987). An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in gram-negative bacteria by marker exchange-eviction mutagenesis. Gene *57*, 239–246.

Rogozin, I.B., Makarova, K.S., Natale, D.A., Spiridonov, A.N., Tatusov, R.L., Wolf, Y.I., Yin, J., and Koonin, E.V. (2002). Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. Nucleic Acids Res. *30*, 4264–4271.

Rosgaard, L., de Porcellinis, A.J., Jacobsen, J.H., Frigaard, N.-U., and Sakuragi, Y. (2012). Bioengineering of carbon fixation, biofuels, and biochemicals in cyanobacteria and plants. J. Biotechnol. *162*, 134–147.

Ruffing, A.M. (2011). Engineered cyanobacteria: teaching an old bug new tricks. Bioeng. Bugs *2*, 136–149.

Ruffing, A.M. (2014). Improved free fatty acid production in cyanobacteria with *Synechococcus* sp. PCC 7002 as Host. Front. Bioeng. Biotechnol. *2*.

Rui, L., Reardon, K.F., and Wood, T.K. (2004). Protein engineering of toluene ortho-monooxygenase of *Burkholderia cepacia* G4 for regiospecific hydroxylation of indole to form various indigoid compounds. Appl. Microbiol. Biotechnol. *66*, 422–429.

Sadre, R., Pfaff, C., and Buchkremer, S. (2012). Plastoquinone-9 biosynthesis in cyanobacteria differs from that in plants and involves a novel 4-hydroxybenzoate solanesyltransferase. Biochem. J. *442*, 621–629.

Saha, R., Verseput, A.T., Berla, B.M., Mueller, T.J., Pakrasi, H.B., and Maranas, C.D. (2012). Reconstruction and comparison of the metabolic potential of cyanobacteria *Cyanothece* sp. ATCC 51142 and *Synechocystis* sp. PCC 6803. PLOS ONE *7*, e48285.

Sakai, M., Ogawa, T., Matsuoka, M., and Fukuda, H. (1997). Photosynthetic conversion of carbon dioxide to ethylene by the recombinant cyanobacterium, *Synechococcus* sp. PCC 7942, which harbors a gene for the ethylene-forming enzyme of *Pseudomonas syringae*. J. Ferment. Bioeng. *84*, 434–443.

Sakai, Y., Abe, K., Nakashima, S., Ellinger, J.J., Ferri, S., Sode, K., and Ikebukuro, K. (2015). Scaffold-fused riboregulators for enhanced gene activation in *Synechocystis* sp. PCC 6803. MicrobiologyOpen *4*, 533–540.

Sakurai, I., Stazic, D., Eisenhut, M., Vuorio, E., Steglich, C., Hess, W.R., and Aro, E.-M. (2012). Positive regulation of *psbA* gene expression by cis-encoded antisense RNAs in *Synechocystis* sp. PCC 6803 [OA]. Plant Physiol. *160*, 1000–1010.

Salis, H.M., Mirsky, E.A., and Voigt, C.A. (2009). Automated design of synthetic ribosome binding sites to precisely control protein expression. Nat. Biotechnol. *27*, 946–950.

Samantaray, S., Bhati, R., and Mallick, N. (2014). Cyanobacterial polyhydroxyalkanoates: an alternative source for plastics. In Cyanobacteria, N.K. Sharma, A.K. Rai, and L.J. Stal, eds. (John Wiley & Sons, Ltd), pp. 227–244.

Savakis, P., Tan, X., Du, W., Branco dos Santos, F., Lu, X., and Hellingwerf, K.J. (2015). Photosynthetic production of glycerol by a recombinant cyanobacterium. J. Biotechnol. *195*, 46–51.

Savakis, P.E., Angermayr, S.A., and Hellingwerf, K.J. (2013). Synthesis of 2,3-butanediol by *Synechocystis* sp. PCC6803 via heterologous expression of a catabolic pathway from lactic acid- and enterobacteria. Metab. Eng. *20*, 121–130.

Scaife, M.A., Nguyen, G.T.D.T., Rico, J., Lambert, D., Helliwell, K.E., and Smith, A.G. (2015). Establishing *Chlamydomonas reinhardtii* as an industrial biotechnology host. Plant J. *82*, 532–546.

Scanlan, D.J., Bloye, S.A., Mann, N.H., Hodgson, D.A., and Carr, N.G. (1990). Construction of *lacZ* promoter probe vectors for use in *Synechococcus*: application to the identification of CO_2-regulated promoters. Gene *90*, 43–49.

Schopf, J.W., and Packer, B.M. (1987). Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. Science *237*, 70–73.

Serino, G., and Maliga, P. (1997). A negative selection scheme based on the expression of cytosine deaminase in plastids. Plant J. *12*, 697–701.

Shao, C.Y., Howe, C.J., Porter, A.J.R., and Glover, L.A. (2002). Novel cyanobacterial biosensor for detection of herbicides. Appl. Environ. Microbiol. *68*, 5026–5033.

Shen, C.R., and Liao, J.C. (2012). Photosynthetic production of 2-methyl-1-butanol from CO_2 in cyanobacterium *Synechococcus elongatus* PCC7942 and characterization of the native acetohydroxyacid synthase. Energy Environ. Sci. *5*, 9574–9583.

Shestakov, S.V., and Khyen, N.T. (1970). Evidence for genetic transformation in blue-green alga *Anacystis nidulans*. Mol. Gen. Genet. MGG *107*, 372–375.

Song, K., Tan, X., Liang, Y., and Lu, X. (2016). The potential of *Synechococcus elongatus* UTEX 2973 for sugar feedstock production. Appl. Microbiol. Biotechnol. *100*, 7865–7875.

Song, Z., Chen, L., Wang, J., Lu, Y., Jiang, W., and Zhang, W. (2014). A transcriptional regulator sll0794 regulates tolerance to biofuel ethanol in photosynthetic *Synechocystis* sp. PCC 6803. Mol. Cell. Proteomics MCP *13*, 3519–3532.

Spence, E., Sarcina, M., Ray, N., Møller, S.G., Mullineaux, C.W., and Robinson, C. (2003). Membrane-specific targeting of green fluorescent protein by the Tat pathway in the cyanobacterium *Synechocystis* PCC6803. Mol. Microbiol. *48*, 1481–1489.

Steinmetz, M., Coq, D.L., Aymerich, S., Gonzy-Tréboul, G., and Gay, P. (1985). The DNA sequence of the gene for the secreted *Bacillus subtilis* enzyme levansucrase and its genetic control sites. Mol. Gen. Genet. MGG *200*, 220–228.

Stoffels, L. (2014). Microalgae as a novel production platform for antibacterial protiens. Ph.D. University College London (University of London).

Stougaard, J. (1993). Substrate-dependent negative selection in plants using a bacterial cytosine deaminase gene. Plant J. *3*, 755–761.

Summers, M.L., Wallis, J.G., Campbell, E.L., and Meeks, J.C. (1995). Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium *Nostoc* sp. strain ATCC 29133. J. Bacteriol. *177*, 6184–6194.

Suzuki, T., Miyake, M., Tokiwa, Y., Saegusa, H., Saito, T., and Asada, Y. (1996). A recombinant cyanobacterium that accumulates poly-(hydroxybutyrate). Biotechnol. Lett. *18*, 1047–1050.

Takahama, K., Matsuoka, M., Nagahama, K., and Ogawa, T. (2003). Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the psbAI locus. J. Biosci. Bioeng. *95*, 302–305.

Takahashi, N., Tamura, M., Kajiwara, S., and Ohtaguchi, K. (1999). Genetically engineered herbicide resistance in cyanobacterium *Synechococcus* sp. by bialaphos resistance gene from *Streptomyces hygroscopicus*. Biotechnol. Lett. *21*, 751–757.

Takeshima, Y., Sugiura, M., and Hagiwara, H. (1994). A novel expression vector for the cyanobacterium, *Synechococcus* PCC 6301. DNA Res. *1*, 181–189.

Tan, X., Yao, L., Gao, Q., Wang, W., Qi, F., and Lu, X. (2011). Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. Metab. Eng. *13*, 169–176.

Tan, X., Liang, F., Cai, K., and Lu, X. (2013). Application of the FLP/FRT recombination system in cyanobacteria for construction of markerless mutants. Appl. Microbiol. Biotechnol. *97*, 6373–6382.

Taton, A., Lis, E., Adin, D.M., Dong, G., Cookson, S., Kay, S.A., Golden, S.S., and Golden, J.W. (2012). Gene transfer in *Leptolyngbya* sp. Strain BL0902, a cyanobacterium suitable for production of biomass and bioproducts. PLOS ONE *7*, e30901.

Taylor, L.A., and Rose, R.E. (1988). A correction in the nucleotide sequence of the Tn903 kanamycin resistance determinant in pUC4K. Nucleic Acids Res. *16*, 358.

Thiel, T., and Poo, H. (1989). Transformation of a filamentous cyanobacterium by electroporation. J. Bacteriol. *171*, 5743–5746.

Thiel, T., Lyons, E.M., Erker, J.C., and Ernst, A. (1995). A second nitrogenase in vegetative cells of a heterocyst-forming cyanobacterium. Proc. Natl. Acad. Sci. *92*, 9358–9362.

Tian, X., Chen, L., Wang, J., Qiao, J., and Zhang, W. (2013). Quantitative proteomics reveals dynamic responses of *Synechocystis* sp. PCC 6803 to next-generation biofuel butanol. J. Proteomics *78*, 326–345.

Tracy, N.I., Chen, D., Crunkleton, D.W., and Price, G.L. (2009). Hydrogenated monoterpenes as diesel fuel additives. Fuel *88*, 2238–2240.

Trautmann, D., Voß, B., Wilde, A., Al-Babili, S., and Hess, W.R. (2012). Microevolution in cyanobacteria: Re-sequencing a motile substrain of *Synechocystis* sp. PCC 6803. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes *19*, 435–448.

Triana, J., Montagud, A., Siurana, M., Fuente, D., Urchueguía, A., Gamermann, D., Torres, J., Tena, J., de Córdoba, P.F., and Urchueguía, J.F. (2014). Generation and evaluation of a genome-scale metabolic network model of *Synechococcus elongatus* PCC7942. Metabolites *4*, 680–698.

Tsinoremas, N.F., Kutach, A.K., Strayer, C.A., and Golden, S.S. (1994). Efficient gene transfer in *Synechococcus* sp. strains PCC 7942 and PCC 6301 by interspecies conjugation and chromosomal recombination. J. Bacteriol. *176*, 6764–6768.

Tyo, K.E.J., Jin, Y.-S., Espinoza, F.A., and Stephanopoulos, G. (2009). Identification of gene disruptions for increased poly-3-hydroxybutyrate accumulation in *Synechocystis* PCC 6803. Biotechnol. Prog. *25*, 1236–1243.

Ungerer, J., Tao, L., Davis, M., Ghirardi, M., Maness, P.-C., and Yu, J. (2012). Sustained photosynthetic conversion of CO_2 to ethylene in recombinant cyanobacterium *Synechocystis* 6803. Energy Environ. Sci. *5*, 8998–9006.

U.S Department of Energy (2002). Steam system opportunity assessment for the pulp and paper, chemical manufacturing, and petroleum refining industries. (Office of Energy Efficency and Renewable Energy).

Varman, A.M., Xiao, Y., Pakrasi, H.B., and Tang, Y.J. (2013). Metabolic engineering of *Synechocystis* sp. strain PCC 6803 for isobutanol production. Appl. Environ. Microbiol. *79*, 908–914.

Varshavsky, A. (1997). The N-end rule pathway of protein degradation. Genes Cells Devoted Mol. Cell. Mech. *2*, 13–28.

Vermaas, W. (1996). Molecular genetics of the cyanobacterium *Synechocystis* sp. PCC 6803: Principles and possible biotechnology applications. J. Appl. Phycol. *8*, 263–273.

Vermaas, W.F. (2001). Photosynthesis and respiration in cyanobacteria. In eLS, (John Wiley & Sons, Ltd), p.

Vermes, A., Guchelaar, H.-J., and Dankert, J. (2000). Flucytosine: a review of its pharmacology, clinical indications, pharmacokinetics, toxicity and drug interactions. J. Antimicrob. Chemother. *46*, 171–179.

Viola, S., Ruhle, T., and Leister, D. (2014). A single vector-based strategy for marker-less gene replacement in *Synechocystis* sp. PCC 6803. Microb. Cell Factories *13*, 4.

Vioque, A. (2007). Transformation of cyanobacteria. In transgenic microalgae as green cell factories, R. León, A. Galván, and E. Fernández, eds. (Springer New York), pp. 12–22.

Voigt, K., Sharma, C.M., Mitschke, J., Joke Lambrecht, S., Voß, B., Hess, W.R., and Steglich, C. (2014). Comparative transcriptomics of two environmentally relevant cyanobacteria reveals unexpected transcriptome diversity. ISME J. *8*, 2056–2068.

Vranová, E., Coman, D., and Gruissem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annu. Rev. Plant Biol. *64*, 665–700.

Wang, B., Kitney, R.I., Joly, N., and Buck, M. (2011). Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. Nat. Commun. *2*, 508.

Wang, B., Wang, J., Zhang, W., and Meldrum, D.R. (2012a). Application of synthetic biology in cyanobacteria and algae. Front. Microbiol. *3*.

Wang, J., Chen, L., Huang, S., Liu, J., Ren, X., Tian, X., Qiao, J., and Zhang, W. (2012b). RNA-seq based identification and mutant validation of gene targets related to ethanol resistance in cyanobacterial *Synechocystis* sp. PCC 6803. Biotechnol. Biofuels *5*, 89.

Watanabe, S., Ohbayashi, R., Kanesaki, Y., Saito, N., Chibazakura, T., Soga, T., and Yoshikawa, H. (2015). Intensive DNA replication and metabolism during the lag phase in cyanobacteria. PLOS ONE *10*, e0136800.

Wendt, K.E., Ungerer, J., Cobb, R.E., Zhao, H., and Pakrasi, H.B. (2016). CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium *Synechococcus elongatus* UTEX 2973. Microb. Cell Factories *15*, 115.

Werner, R., and Mergenhagen, D. Mating type determination of *Chlamydomonas reinhardtii* by PCR. Plant Mol. Biol. Report. *16*, 295–299.

Williams, J.G.K. (1988). [85] Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in *Synechocystis* 6803. B.-M. in Enzymology, ed. (Academic Press), pp. 766–778.

Williams, D.C., McGarvey, D.J., Katahira, E.J., and Croteau, R. (1998). Truncation of limonene synthase preprotein provides a fully active "pseudomature" form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. Biochemistry (Mosc.) *37*, 12213–12220.

Willrodt, C., David, C., Cornelissen, S., Bühler, B., Julsing, M.K., and Schmid, A. (2014). Engineering the productivity of recombinant *Escherichia coli* for limonene formation from glycerol in minimal media. Biotechnol. J. *9*, 1000–1012.

Wolk, C.P., Vonshak, A., Kehoe, P., and Elhai, J. (1984). Construction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria. Proc. Natl. Acad. Sci. U. S. A. *81*, 1561–1565.

Wubbolts, M.G., Reuvekamp, P., and Witholt, B. (1994). TOL plasmid-specified xylene oxygenase is a wide substrate range monooxygenase capable of olefin epoxidation. Enzyme Microb. Technol. *16*, 608–615.

Xu, Y., and Green, B.D. (2012). Compositions and methods to remove genetic markers using counter-selection.

Xu, H., Vavilin, D., Funk, C., and Vermaas, W. (2004). Multiple deletions of small cab-like proteins in the cyanobacterium *Synechocystis* sp. PCC 6803 consequences for pigment biosynthesis and accumulation. J. Biol. Chem. *279*, 27971–27979.

Yao, L., Cengic, I., Anfelt, J., and Hudson, E.P. (2016). Multiple gene repression in cyanobacteria using CRISPRi. ACS Synth. Biol. *5*, 207–212.

Yoshihara, S., Geng, X., Okamoto, S., Yura, K., Murata, T., Go, M., Ohmori, M., and Ikeuchi, M. (2001). Mutational analysis of genes involved in pilus structure, motilityand transformation competency in the unicellular motile cyanobacterium *Synechocystis* sp. PCC6803. Plant Cell Physiol. *42*, 63–73.

Young, R.E.B., and Purton, S. (2014). Cytosine deaminase as a negative selectable marker for the microalgal chloroplast: a strategy for the isolation of nuclear mutations that affect chloroplast gene expression. Plant J. *80*, 915–925.

Yu, J., Liberton, M., Cliften, P.F., Head, R.D., Jacobs, J.M., Smith, R.D., Koppenaal, D.W., Brand, J.J., and Pakrasi, H.B. (2015). *Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO2. Sci. Rep. *5*, 8132.

Yu, Y., You, L., Liu, D., Hollinshead, W., Tang, Y.J., and Zhang, F. (2013). Development of *Synechocystis* sp. PCC 6803 as a Phototrophic Cell Factory. Mar. Drugs *11*, 2894–2916.

Yuba, A., Yazaki, K., Tabata, M., Honda, G., and Croteau, R. (1996). cDNA cloning, characterization, and functional expression of 4S-(-)-limonene synthase from *Perilla frutescens*. Arch. Biochem. Biophys. *332*, 280–287.

Zang, X., Liu, B., Liu, S., Arunakumara, K.K.I.U., and Zhang, X. (2007). Optimum conditions for transformation of *Synechocystis* sp. PCC 6803. J. Microbiol. Seoul Korea *45*, 241–245.

Zerulla, K., Ludt, K., and Soppa, J. (2016). The ploidy level of *Synechocystis* sp. PCC 6803 is highly variable and is influenced by growth phase and by chemical and physical external parameters. Microbiology *162*, 730–739.

Zess, E.K., Begemann, M.B., and Pfleger, B.F. (2016). Construction of new synthetic biology tools for the control of gene expression in the cyanobacterium *Synechococcus* sp. strain PCC 7002. Biotechnol. Bioeng. *113*, 424–432.

Zhang, S., Liu, Y., and Bryant, D.A. (2015). Metabolic engineering of *Synechococcus* sp. PCC 7002 to produce poly-3-hydroxybutyrate and poly-3-hydroxybutyrate-co-4-hydroxybutyrate. Metab. Eng. *32*, 174–183.

Zhang, Y., Pu, H., Wang, Q., Cheng, S., Zhao, W., Zhang, Y., and Zhao, J. (2007). PII is important in regulation of nitrogen metabolism but not required for heterocyst formation in the cyanobacterium *Anabaena* sp. PCC 7120. J. Biol. Chem. *282*, 33641–33648.

Zhao, C., Li, Z., Li, T., Zhang, Y., Bryant, D.A., and Zhao, J. (2015). High-yield production of extracellular type-I cellulose by the cyanobacterium *Synechococcus* sp. PCC 7002. Cell Discov. *1*, 15004.

Zhou, J., Zhang, H., Zhang, Y., Li, Y., and Ma, Y. (2012). Designing and creating a modularized synthetic pathway in cyanobacterium *Synechocystis* enables production of acetone from carbon dioxide. Metab. Eng. *14*, 394–400.

Zhou, J., Zhang, H., Meng, H., Zhu, Y., Bao, G., Zhang, Y., Li, Y., and Ma, Y. (2014). Discovery of a super-strong promoter enables efficient production of heterologous proteins in cyanobacteria. Sci. Rep. *4*.

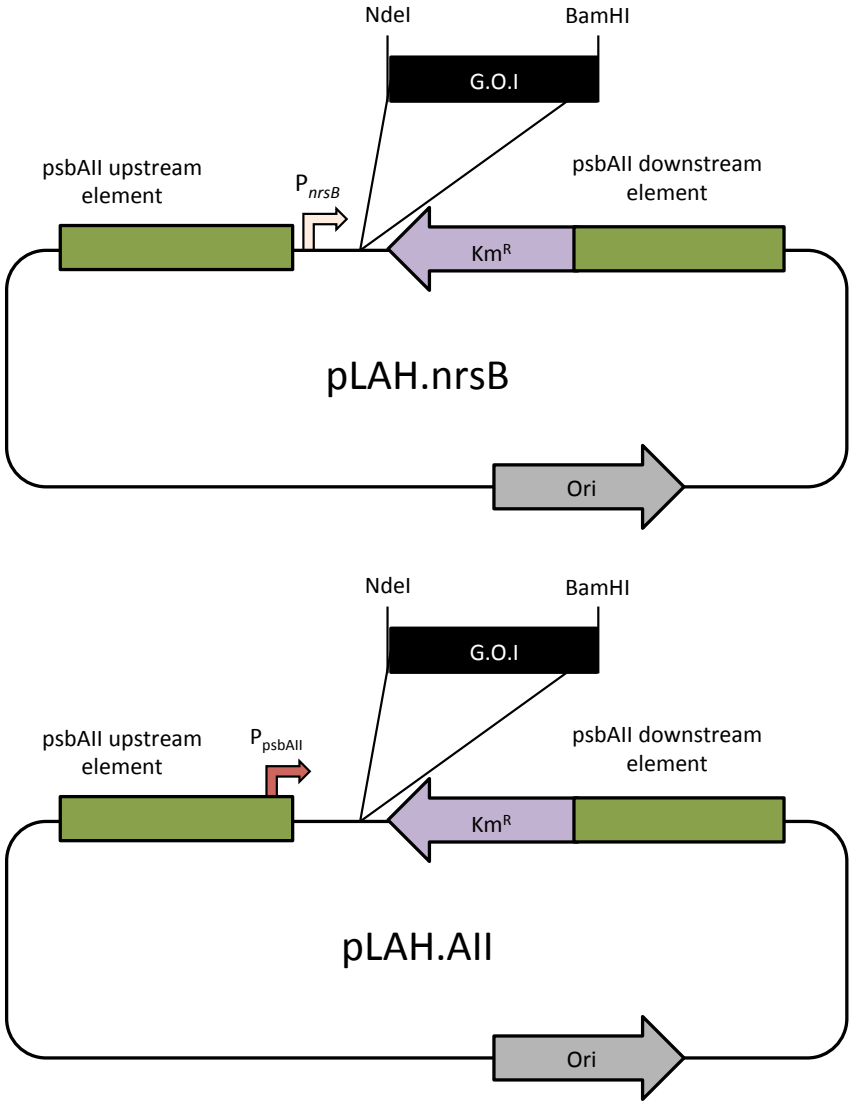
Zhu, H., Ren, X., Wang, J., Song, Z., Shi, M., Qiao, J., Tian, X., Liu, J., Chen, L., and Zhang, W. (2013). Integrated OMICS guided engineering of biofuel butanol-tolerance in photosynthetic *Synechocystis* sp. PCC 6803. Biotechnol. Biofuels *6*, 106.

Appendix 1: Primer List

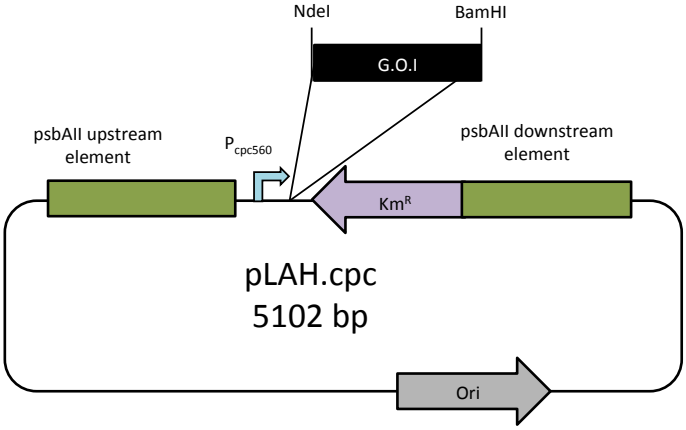
Primer name	Sequence (5' to 3')
6803_codA_F	CGC TCT CAT ATG TCT AAC AAC GC
6803_codA_R	GCA TGC GGA TCC TTA TTA AGC G
6803_sacB_F	TCT TCA AAC AGG AGG GCT GG
6803_sacB_R	CGG TTA GCC ATT TGC CTG C
AII.LS_F	ATG CTA GCG CTT CGT GTA TAT TAA CTT CC
AII.LS_R	TCG TTA ACA GGT CGG GGA TCC TTA GG
AL.KanR.4op.R	CGG GGT ACC TCA ATT CAA TTC ATC AAG
AL.ZmR.4op.F	TAT AAC CAT ATG GCC AAG TTG ACC AGT GCC
Ben seq ups F	CCC AGA ACT ATG GTA AAG GCG
Ble_Fw	GGC ATT AAT GGC CAA GTT GAC C
Ble_IG0bp_Re	AGT TCT TCT CCT TTA CTC ATT CAG TCC TGC TCC TCG GCC ACG
Ble_IG1bp_Re	AGT TCT TCT CCT TTA CTC ATC TCA GTC CTG CTC CTC GGC CAC G
Ble_IG2_Re	GGC AAA ATC CGG GGG GCT GGT CAA CTA GTA ATT ATC AGT CCT GCT CCT CGG
Ble_IG4_Re	TCG ATT ACC TCC ACT ATT GCC ATC AGT CCT GCT CCT CGG
Ble_IG5_Re	TTT ATT CCT TTG CCT TTT AGT TCA GTC CTG CTC CTC GG
C.FragBR	ATA AAC CGC CTG ATA GGT GTT GG
codAmidF	TCG TTG AAA CAG TAG CTG C
FO.kmF	CCT TTT AAC AGC GAT CGC G
FO.kmR	GGA ATC GAA TGC AAC CGG
GFP_0bp_F	CGA TGT TCC AGA TTA CGC TTA AAT GAG TAA AGG AGA AGA AC
GFP_1bp_F	CGA TGT TCC AGA TTA CGC TTA ACA TGA GTA AAG GAG AAG AAC
GFP_IG0bp_Fw	GGC CGA GGA GCA GGA CTG AAT GAG TAA AGG AGA AGA AC
GFP_IG1bp_Fw	GGC CGA GGA GCA GGA CTG AGA TGA GTA AAG GAG AAG AAC
GFP_IG2_Fw	TAA TTA CTA GTT GAC CAG CCC CCC GGA TTT TGC CAT GAG TAA AGG AGA AGA AC
GFP_IG4_Fw	GGC AAT AGT GGA GGT AAT CGA ATG AGT AAA GGA GAA GAA C
GFP_IG5_Fw	CTA AAA GGC AAA GGA ATA AAA CTA TGA GTA AAG GAG AAG AAC
GFP_Re	GGA GAT CTT ATT TGT AGA GCT CAT CC
IR706.Ex F	ACC AAG AGC AAC AAA GAC AGG
IR706.Ex R	CTT ATC AAC AGG GAT TCC GG
IR706.Ex.Flank.F	CAT CCG AAT CTT TAA ACC C
IR706.Ex.Flank.R	GAA CCT TGG ATT TCA TCA GG
IR706F	TCA GCG TAC AGA TCA TTC GG
IR706R	GCT ATG TAC ACT GGC TAA GG
KmRSacBRF	TTT GCT AGC AAT TCG AGC TCG GTA CCC G
KmRSacBRR	AAA GCT AGC CCC ATC ACA TAT ACC TGC C
LSmidF	TGC CCG TGA ATT TGC CAC C

Primer name	Sequence (5’ to 3’)
NptI_outF	GGC CTC AAC ACG ATT TTA CG
nrsBseqF	CCG TCT CAT CTT CCA CCA GC
Out.GFP.R	AAC ACC TTC TTC ACG AGG CAG ACC
Pcpc_ins_Fw	TGG ACT CCC TCA GTT TAT CCG
pJET F	TTT TAA CTT GGA GCA GGT TCC
pJET R	GTT TTC ATG AGA GTC GAT TGC
pLAH_codAF	CGT GGC TAG CAA CTT CCT GTT ACA AAG C
pLAH_codAR	TAC AGC TAG CGT TAT GAG CCA TAT TCA ACG
pLAH.nrsBR.F	ATG CGA CTA CGG GCA AAG AGG
pLAH.pcpc.F	GCA CCT GTA GAG AAG AGT CC
pLAH.pro.R	TCG GGG ATC CTC TAG AGT CG
Slr0611 F	AAA TGA AAA CCA TTG CCC CGC C
Slr0611 R	CAA TCG TAG GGT AGG GAA GC
splB_0bp_R	GTT CTT CTC CTT TAC TCA TTT AAG CGT AAT CTG GAA CAT CG
splB_1bp_R	GTT CTT CTC CTT TAC TCA TCT TAA GCG TAA TCT GGA ACA TCG
splB_F	GCA TTA ATG AAC AAA AAC GTA GTT ATT AAA TCA TTA GC
StyOx_End_F	GTG TAG TCA ATT TGC CCC CAC C
StyOx_seq1_R	CCG GGT TGG CTA AGT ATT GC
StyOx_seq2_F	TTC TAG CAA CCC CTC CCT CG
StyOx_seq3_R	CCA AAG CAC AGT GCA AAG G
StyOx_seq4_F	TAA ACC CCA ACA CAT TCG CC
StyOx_seq5_F	AAC CCC AGC GAG CCT TAT GC
StyOx_seq6_R	AAA CTC ATG GTG ACA GCA CG

Appendix 2: Plasmids used in this study



Plasmids pLAH.nrsB and pLAH.AII designed and created by Lamya Al-Haj (Al-Haj, 2014). The NdeI and BamHI sites are present for inserting the G.O.I.



Plasmid pLAH.cpc created by Aaron Lau , again the NdeI and BamHI sites are present for inserting the G.O.I.

*Synthesised by GeneArt

Plasmid name	**Backbone**	**Insert**
ls_pMA-RQ*	pMA-TRQ	*ls* with a HA-tag sequence at the SfiI restriction sites.
pLAH.nrsB.LS	pLAH.nrsB	*ls* with a HA-tag sequence at the NdeI and BamHI restriction sites.
pLAH.AII.LS	pLAH.AII	*ls* with a HA-tag sequence at the NdeI and BamHI restriction sites.
pLAH.cpc.LS	pLAH.cpc	*ls* with a HA-tag sequence at the NdeI and BamHI restriction sites.
pJET.sds	pJET1.2/blunt	*sds* (*sllr0611*) gene amplified from genomic WT *Synechocystis* at the multiple cloning sites (MCS).
pJET.sds::Km 3	pJET.sds	Kanamycin resistant cassette from pUC4K at the HincII and SmaI site.
pJET.sds::Km 6	pJET.sds	Kanamycin resistant cassette from pUC4K at the HincII and SmaI site. Differs from pJET.sds::Km 6 as kanamycin resistant cassette is in the opposite orientation.
pLAH.nrsB_cat	pLAH.nrsB	*cat* cassette from pARG1.3 inserted at the DraI restriction site.
pJET.zsk	pJET1.2/blunt	*Ble*-IG1-*aadA*-IG2-*aphA-6* amplified from 4op at the MCS.
pLAH.nrsB.zsk	pLAH.nrsB_cat	*Ble*-IG1-*aadA*-IG2-*aphA-6* from pJET.zsk ligated at the NdeI and Acc65I site.
pLAH.ble.IG2.GFP	pLAH.cpc	*ble*-IG2-*gfpuv* PCR product is ligated at the NdeI and BamHI site.
pLAH.ble.IG4.GFP	pLAH.cpc	*ble*-IG4-*gfpuv* PCR product is ligated at the NdeI and BamHI site.
pLAH.ble.IG5.GFP	pLAH.cpc	*ble*-IG5-*gfpuv* PCR product is ligated at the NdeI and BamHI site.
pLAH.ble.IG0bp.GFP	pLAH.cpc	*ble*-IG0bp-*gfpuv* PCR product is ligated

Plasmid name	Backbone	Insert
		at the NdeI and BamHI site.
pLAH.ble.IG1bp.GFP	pLAH.cpc	*ble*-IG1bp-*gfpuv* PCR product is ligated at the NdeI and BamHI site.
pLAH.splB.IG0bp.GFP	pLAH.cpc	*splB*-IG0bp-*gfpuv* PCR product is ligated at the NdeI and BamHI site.
pLAH.splB.IG1bp.GFP	pLAH.cpc	*splB*-IG1bp-*gfpuv* PCR product is ligated at the NdeI and BamHI site.
pIR706	pJET1.2/blunt	IR706, neutral site, amplified from the genomic WT *Synechocystis* at the MCS
pIR706.cat	pIR706	*cat* cassette from pARG1.3 inserted at the HpaI site.
pJET.sacB.nptI	pJET1.2/blunt	*nptI* and *sacB* from pUM24Cm inserted at the MCS.
pJET.sacB.nptI'	pJET.sacB.nptI	DNA between AfeI and EcoRV digested out.
pIR706.sacB.nptI	pIR706	*nptI* and *sacB* from pJET.sacB.nptI inserted at NheI restriction site.
pIR706.sacB.nptI'	pIR706	*nptI* and *sacB* from pJET.sacB.nptI' inserted at NheI restriction site.
pLAH.AII.codA	pLAH.AII	*codA* with a HA-tag sequence amplified from pRY127d inserted at the NdeI and BamHI restriction sites.
pIR706.codA	pIR706	*codA* with a HA-tag sequence under the control of P_{psbAII} inserted and the kanamycin resistance cassette from pLAH.AII.codA at the NheI site.
pIR706.Ex	pJET1.2/blunt	Extended neutral site, IR706.Ex, amplified from the genomic WT *Synechocystis* at the MCS
pIR706.Ex.cat	pIR706.Ex	*cat* cassette from pARG1.3 inserted at the HpaI site.
pIR706.Ex.LS	pIR706.Ex	*ls* with a HA-tag sequence under the control of P_{psbAII} from pLAH.AII.LS inserted at the NheI and HpaI site.
pIR706.Ex.cpc	pIR706.Ex	P_{cpc560} and the NdeI and BamHI restriction sites from pLAH.cpc inserted at the HpaI site.
pIR706.Ex.nrsB	pIR706.Ex	P_{nrsB} and the NdeI and BamHI restriction sites from pLAH.cpc inserted at the HpaI site.
SO_Part1_pMA-T*	pMA-T	SO_Part1 containing *PAL2*-IG4-*FDC1* up to SphI RE site.
SO_Part2_pMA-T*	pMA-T	SO_Part2 containing *FDC1* after SphI-IG5-*styA*-IG4-*styB*.
pIR706.Ex.nrsB.SO1	pIR706.Ex.nrsB	SO_Part1 from SO_Part1_pMA-T inserted at the NdaI and BamHI restriction sites.
pIR706.Ex.cpc.SO1	pIR706.Ex.cpc	SO_Part1 from SO_Part1_pMA-T inserted at the NdeI and BamHI restriction sites.
pIR706.Ex.nrsB.SO	pIR706.Ex.nrsB.SO1	SO_Part2 from SO_Part2_pMA-T inserted at the SphI and BamHI restriction sites.
pIR706.Ex.cpc.SO	pIR706.Ex.cpc	SO operon from pIR706.Ex.nrsB.SO inserted into NdeI and BamHI restriction

Plasmid name	Backbone	Insert
		sites.
pLAH.cpc.SO	pLAH.cpc	SO operon from pIR706.Ex.nrsB.SO inserted into NdeI and BamHI restriction sites.
pLAH.nrsB.SO	pLAH.nrsB	SO operon from pIR706.Ex.nrsB.SO inserted into NdeI and BamHI restriction sites.

Appendix 3: Transgenic strains created in this study

Strain name (in this study)	Description
6803.nrsB.LS	The *psbAII* promoter and gene is replaced by *ls* with a HA-tag sequence under the control of the P_{nrsB} and a kanamycin resistance cassette.
6803.AII.LS	The *psbAII* gene is replaced by *ls* with a HA-tag sequence under the control of the P_{psbAII} and a kanamycin resistance cassette.
6803.cpc.LS	The *psbAII* promoter and gene is replaced by *ls* with a HA-tag sequence under the control of the P_{cpc560} and a kanamycin resistance cassette.
6803.nrsB.zsk	The *psbAII* promoter and gene is replaced by *ble, aadA* and *aphA-6* under the control of the P_{nrsB}. Resistant to zeocin, spectinomycin and kanamycin
6803.ble.IG2.GFP	The *psbAII* promoter and gene is replaced by *ble* and *gfp* under the control of the P_{cpc560} and a kanamycin resistance cassette. Between ble and GFP is IG2. Also resistant to zeocin.
6803.ble.IG4.GFP	The *psbAII* promoter and gene is replaced by *ble* and *gfp* under the control of the P_{cpc560} and a kanamycin resistance cassette. Between ble and GFP is IG4. Also resistant to zeocin.
6803.ble.IG5.GFP	The *psbAII* promoter and gene is replaced by *ble* and *gfp* under the control of the P_{cpc560} and a kanamycin resistance cassette. Between ble and GFP is IG5. Also resistant to zeocin.
6803.ble.0bp.GFP	The *psbAII* promoter and gene is replaced by *ble* and *gfp* under the control of the P_{cpc560} and a kanamycin resistance cassette. Between ble and GFP is IG0bp. Also resistant to zeocin.
6803.ble.1bp.GFP	The *psbAII* promoter and gene is replaced by *ble* and *gfp* under the control of the P_{cpc560} and a kanamycin resistance cassette. Between ble and GFP is IG1bp. Also resistant to zeocin.
6803.splB.0bp.GFP	The *psbAII* promoter and gene is replaced by *splB* and *gfp* under the control of the P_{cpc560} and a kanamycin resistance cassette. Between *splB* and GFP is IG0bp. Also resistant to zeocin.
6803.splB.1bp.GFP	The *psbAII* promoter and gene is replaced by *splB* and *gfp* under the control of the P_{cpc560} and a kanamycin resistance cassette. Between *splB* and GFP is IG1bp. Also resistant to zeocin.
6803.IR706.sacB	The selected neutral region between *slr1340* and *sll1255* at the NheI site has *sacB* and *nptI* expressed under their own promoter. Resistant to kanamycin.
6803.IR706.sacB'	The selected neutral region between *slr1340* and *sll1255* at the NheI site has *sacB* and *nptI* expressed under their own promoter. Differs from 6803.IR706.sacB as a predicted transposon was removed between *sacB* and *nptI*. Resistant to kanamycin.
6803.AII.codA	The *psbAII* gene is replaced by *codA* with a HA-tagged sequence under the control of P_{psbAII} and a kanamycin resistance cassette.
6803.IR706.codA	The selected neutral region between *slr1340* and *sll1255* at the NheI site has *codA* under the control of the *psbAII* and a kanamycin resistance cassette.
6803.IR706.cat	The selected neutral region between *slr1340* and *sll1255* at the HpaI site has *cat* under the control of its own promoter. Resistance

Strain name (in this study)	Description
	to chloramphenicol.
6803.IR706.LS	Replacing the region between the NheI and HpaI site, found within the neutral region *slr1340* and *sll1255*, is *ls* with a HA-tag sequence under the control of the P_{psbAII}. Marker-less strains.
6803.AII.cpc.SO	The *psbAII* promoter and gene is replaced by the *PAL2*, *FDC1, styA* and *styB* under the control of the P_{cpc560} and a kanamycin resistance cassette.
6803.AII.nrsB.SO	The *psbAII* promoter and gene is replaced by the *PAL2*, *FDC1, styA* and *styB* under the control of the P_{nrsB} and a kanamycin resistance cassette.
6803.IR706.cpc	The selected neutral region between *slr1340* and *sll1255* at the HpaI site has no gene-of-interest under the control of P_{cpc560}.

Appendix 4: DNA Constructs

Synthetic LS gene sequence

Genes:
ls
HA epitope tag
TAA – Stop codon

Restriction sites:
NdeI
BamHI

CATATGACCGAACGGCGTTCCGGCAATTACAATCCCTCCCGGTGGGATGTGAACTTTATTCAAT
CCTTGTTGTCCGATTACAAAGAAGATAAACACGTTATTCGTGCCTCCGAATTGGTGACCTTGGT
GAAAATGGAATTGGAAAAAGAAACCGATCAAATTCGGCAGCTGGAATTGATTGATGATTTGCAA
CGGATGGGCTTGTCCGATCATTTTCAGAACGAATTTAAAGAAATTTTGTCCTCCATTTATTTGG
ATCATCATTATTACAAAAATCCCTTTCCCAAAGAAGAACGGGATTTGTATTCCACCTCCTTGGC
CTTTCGGTTGTTGCGGGAACATGGCTTTCAAGTGGCCCAAGAAGTGTTTGATTCCTTTAAAAAC
GAAGAAGGGGAATTTAAAGAATCCCTGTCCGATGATACCCGTGGCTTGTTGCAATTGTATGAAG
CCTCCTTTTTGTTGACCGAAGGCGAAACCACCTTGGAAAGTGCCCGTGAATTTGCCACCAAATT
TTTGGAAGAAAAAGTGAATGAAGGCGGTGTGGATGGCGATTTGTTGACCCGGATTGCCTATTCC
TTGGATATTCCCTTGCATTGGCGGATTAAACGTCCCAATGCTCCCGTGTGGATTGAATGGTATC
GGAAACGTCCCGATATGAATCCCGTGGTGTTGGAATTGGCCATTTTGGATTTGAATATTGTGCA
GGCCCAGTTTCAAGAAGAATTGAAAGAATCCTTTCGGTGGTGGCGGAATACCGGCTTTGTGGAA
AAATTGCCCTTTGCCCGTGATCGGTTGGTGGAATGTTATTTTTGGAATACCGGCATTATTGAAC
CCCGTCAACATGCCAGTGCCCGTATTATGATGGGCAAAGTGAATGCCTTGATTACCGTGATTGA
TGACATTTACGATGTGTATGGCACCTTGGAAGAACTGGAACAATTTACCGATTTGATTCGGCGT
TGGGACATTAATTCCATTGATCAATTGCCCGATTACATGCAATTGTGTTTTTTGGCCCTGAACA
ACTTTGTGGATGATACCTCCTATGACGTGATGAAAGAAAAAGGCGTGAATGTGATTCCCTATTT
GCGGCAATCCTGGGTGGATTTGGCCGATAAATACATGGTGGAAGCCCGTTGGTTTTATGGCGGT
CATAAACCCAGTTTGGAAGAATATTTGGAAAATTCCTGGCAGTCCATTTCCGGTCCCTGTATGT
TGACCCATATTTTTTTTCGGGTGACCGACTCCTTTACCAAAGAAACCGTGGATTCCTTGTACAA
ATATCATGATTTGGTGCGGTGGTCCTCCTTTGTGTTGCGGTTGGCCGATGATTTGGGCACCTCC
GTGGAAGAAGTGAGTCGGGGTGATGTGCCCAAATCCTTGCAATGTTATATGTCCGATTATAATG
CCTCCGAAGCCGAAGCCCGTAAACATGTGAAATGGTTGATTGCCGAAGTGTGGAAAAAAATGAA
TGCCGAACGGGTGTCCAAAGATTCCCCCTTTGGCAAAGATTTTATTGGCTGTGCCGTGGATTTG
GGTCGGATGGCCCAATTGATGTATCATAATGGTGATGGCCATGGCACCCAACATCCCATTATTC
ATCAACAAATGACCCGGACCTTGTTTGAACCCTTTGCCTATCCCTATGATGTGCCCGATTATGC
CTAAGGATCC

Synthetic SO operon gene sequence

Genes:
PAL2
FDC1
StyA
StyB
TAA – Stop codon

Intergenic Regions:
IG4
IG5

Restriction sites:
NdeI
BamHI
SacI

CATATGGACCAAATCGAGGCCATGTTGTGTGGGGGGGGCGAAAAGACCAAAGTAGCGGTCACCA
CCAAGACTCTCGCGGACCCTTTGAACTGGGGGTTGGCCGCAGATCAGATGAAGGGCAGCCACTT
AGATGAGGTTAAAAAAATGGTGGAGGAATATCGCCGGCCTGTAGTAAACCTCGGCGGAGAAACC
CTGACTATTGGACAAGTGGCAGCAATTAGCACGGTGGGCGGTAGTGTGAAAGTGGAGCTAGCCG
AAACCTCCCGAGCCGGTGTGAAGGCTAGCTCCGACTGGGTGATGGAAAGTATGAATAAGGGAAC
CGATAGCTATGGTGTAACAACCGGCTTCGGGGCGACCAGTCACCGGCGAACCAAGAATGGCACC
GCTCTCCAAACGGAATTGATCCGCTTCTTAAATGCTGGAATTTTTGGGAACACCAAAGAAACCT
GCCATACACTGCCTCAATCTGCAACCCGGGCGGCCATGCTCGTGCGAGTCAACACTCTACTGCA
GGGTTACAGTGGCATTCGGTTTGAGATTTTAGAAGCCATCACTTCCTTGCTCAATCACAACATC
TCTCCCAGTCTCCCGCTCCGCGGAACTATCACTGCCTCCGGCGATCTGGTTCCCCTGTCCTATA
TTGCTGGCCTGCTAACCGGTCGTCCAAATTCTAAAGCCACGGGCCCTGATGGCGAAAGTTTGAC
AGCCAAAGAAGCATTTGAGAAAGCGGGTATTAGTACAGGTTTCTTTGATTTGCAGCCGAAGGAG
GGCCTGGCCCTCGTGAACGGCACCGCGGTTGGCAGTGGGATGGCGTCTATGGTTCTATTTGAAG
CCAATGTGCAAGCGGTGCTCGCGGAAGTTTTATCCGCGATTTTTGCCGAAGTGATGTCCGGTAA
ACCCGAATTCACAGACCATTTAACCCACCGATTGAAACATCATCCGGGCCAGATTGAAGCCGCC
GCCATTATGGAACACATTCTAGATGGTAGCAGTTATATGAAGCTAGCCCAAAAGGTCCATGAGA
TGGCTCCCCTGCAGAAACCGAAACAAGATCGCTATGCTCTACGTACCTCCCCCCAGTGGTTAGG
TCCGCAAATCGAGGTGATCCGGCAAGCCACAAAAAGTATCGAACGCGAAATCAATTCTGTTAAC
GATAATCCCTTGATCGACGTGTCCCGTAATAAAGCCATCCACGGTGGTAATTTTCAAGGGACTC
CCATTGGGGTGTCTATGGACAACACCCGGCTCGCCATTGCCGCGATTGGCAAGTTAATGTTCGC
CCAATTTTCCGAATTAGTCAATGATTTTTACAACAACGGCCTCCCATCTAATTTAACTGCTTCT
AGTAACCCATCCCTGGATTACGGCTTCAAGGGCGCCGAGATTGCCATGGCCTCCTATTGTAGCG
AACTACAATATCTGGCCAATCCTGTGACTTCCCATGTTCAATCCGCGGAACAACATAACCAGGA
TGTGAATAGCCTGGGTTTAATCAGTAGTCGCAAGACATCTGAAGCAGTTGACATTCTAAAATTA
ATGAGCACAACATTCCTGGTAGGAATCTGTCAGGCTGTAGATCTCCGGCATTTGGAAGAAAATT
TACGTCAAACCGTGAAAAATACCGTTAGTCAGGTAGCGAAGAAAGTATTGACCACCGGGATTAA
TGGCGAGTTGCATCCCTCTCGATTTTGTGAAAAAGACCTGCTGAAGGTGGTAGACCGGGAACAG
GTTTTTACATACGTAGACGATCCCTGTTCCGCTACATACCCACTCATGCAACGGTTGCGCCAAG
TGATTGTTGATCATGCGCTGTCCAATGGTGAAACTGAGAAGAACGCTGTTACTTCTATTTTCCA
AAAAATTGGGGCCTTTGAGGAGGAACTGAAAGCCGTATTGCCTAAAGAAGTTGAAGCTGCTCGC
GCCGCGTATGGTAATGGTACGGCCCCCATTCCCAATCGCATTAAAGAATGTCGGTCTTATCCCC
TGTACCGGTTTGTCCGCGAAGAACTAGGAACGAAATTATTAACCGGTGAGAAGGTGGTCTCCCC
TGGCGAAGAATTTGATAAAGTCTTTACCGCCATGTGCGAAGGCAAACTCATTGATCCGTTGATG
GACTGTCTAAAGGAATGGAACGGAGCACCTATTCCTATTTGTTAATGGCAATAGTGGAGGTAAT
CGAATGCGAAAGTTGAACCCTGCACTAGAATTTCGGGATTTTATCCAAGTTCTGAAAGATGAGG
ATGATTTGATCGAAATTACGGAAGAAATTGACCCCAATTTGGAAGTCGGTGCCATTATGCGAAA
AGCCTACGAGAGCCACTTGCCAGCTCCATTGTTTAAGAATCTAAAAGGTGCCAGCAAGGACCTA
TTCTCTATTTTAGGCTGCCCCGCCGGTCTACGCTCTAAAGAAAAAGGGGATCATGGTCGGATTG
CCCATCACTTAGGTCTGGACCCTAAAACCACAATTAAGGAAATTATTGACTATCTGTTGGAATG
CAAAGAAAAAGAACCGTTGCCGCCTATCACAGTGCCCGTCAGCAGTGCCCCGTGCAAAACCCAT
ATTTTAAGTGAAGAAAAAATTCACCTGCAATCTCTACCGACCCCGTATTTGCACGTTTCCGATG
GGGGGAAATACTTACAAACCTATGGCATGTGGATTCTACAGACTCCCGATAAAAAATGGACCAA
TTGGTCTATTGCTCGGGGTATGGTCGTGGATGATAAACATATTACGGGTTTAGTAATTAAACCT
CAACACATCCGCCAGATTGCCGATTCCTGGGCCGCAATTGGTAAAGCTAATGAAATTCCCTTTG
CTCTATGTTTTGGTGTGCCACCCGCTGCCATTTTGGTGAGTTCCATGCCCATCCCGGAAGGTGT
ATCCGAGTCCGATTACGTAGGGGCTATCTTGGGTGAATCCGTGCCAGTTGTGAAGTGTGAGACC
AACGATTTAATGGTCCCGGCTACTAGTGAGATGGTCTTTGAAGGGACCCTGTCCTTGACAGATA
CGCACTTAGAAGGGCCGTTTGGTGAAATGCACGGCTACGTCTTTAAGTCCCAAGGTCATCCTTG
TCCCTTATATACCGTTAAAGCTATGTCTTATCGCGACAATGCGATCTTGCCCGTGAGTAATCCT
GGTTTATGCACTGACGAGACGCATACTCTGATCGGCTCCCTCGTTGCCACCGAAGCTAAGGAGC
TCGCAATTGAAAGCGGCCTCCCGATCTTGGATGCCTTTATGCCGTATGAAGCGCAAGCCTTGTG
GCTGATCCTCAAAGTCGATTTAAAAGGCCTGCAAGCCTTGAAAACGACACCAGAGGAGTTTTGT

AAGAAAGTGGGTGATATCTACTTCCGTACTAAGGTTGGGTTTATTGTTCATGAAATTATTCTAG
TCGCCGATGATATTGACATTTTCAATTTTAAAGAGGTTATCTGGGCTTACGTGACGCGGCACAC
TCCCGTAGCCGATCAAATGGCTTTTGACGATGTGACTAGTTTTCCTCTCGCCCCCTTTGTGTCT
CAAAGCTCCCGTAGTAAGACTATGAAAGGCGGCAAATGTGTGACCAATTGCATTTTTCGTCAAC
AATACGAACGGTCCTTTGACTATATCACGTGTAACTTCGAGAAAGGTTATCCTAAGGGTCTGGT
GGATAAAGTGAATGAAAACTGGAAACGGTACGGGTACAAGTAAACTAAAAGGCAAAGGAATAAA
ACTATGAAAAAGCGCATCGGCATCGTGGGAGCGGGAACCGCCGGCTTGCATTTGGGCTTGTTCC
TGCGTCAGCATGATGTGGATGTCACGGTGTATACCGATCGTAAACCCGATGAATACAGCGGCTT
GCGGCTATTAAATACCGTAGCGCACAATGCCGTTACAGTCCAGCGAGAAGTCGCTCTAGATGTT
AACGAATGGCCTTCCGAAGAATTTGGCTATTTTGGACATTATTACTATGTGGGTGGGCCACAGC
CTATGCGCTTTTATGGCGATTTGAAAGCCCCCAGTCGAGCAGTGGATTATCGCTTGTACCAGCC
CATGCTGATGCGGGCACTGGAGGCTCGCGGTGGAAAATTTTGCTATGATGCTGTCAGCGCGGAA
GACTTAGAAGGCCTGTCTGAGCAATACGATTTACTCGTAGTGTGCACCGGGAAATACGCTCTGG
GCAAAGTCTTTGAGAAGCAGTCCGAGAACAGTCCGTTTGAGAAGCCTCAGCGTGCTCTATGTGT
TGGATTATTTAAAGGCATTAAAGAAGCCCCGATTCGAGCCGTGACCATGTCTTTTAGTCCGGGG
CATGGAGAACTAATCGAAATTCCCACATTAAGCTTTAATGGAATGAGTACCGCGCTAGTTTTGG
AAAATCATATTGGCTCTGACCTAGAGGTCCTGGCCCACACAAAATACGACGATGACCCGCGCGC
CTTTCTAGATTTGATGTTAGAAAAACTAGGAAAACATCATCCCTCTGTCGCCGAACGCATTGAT
CCTGCCGAATTCGATCTAGCTAATAGCTCCTTGGATATCTTGCAGGGAGGAGTTGTGCCGGCCT
TTCGAGATGGGCACGCGACCTTGAACAACGGTAAAACTATTATCGGCCTAGGTGATATTCAAGC
CACCGTCGATCCGGTTCTGGGGCAAGGCGCGAATATGGCGTCTTATGCAGCCTGGATTTTAGGG
GAAGAAATTCTAGCGCATTCTGTATATGATCTACGCTTTAGTGAACATTTAGAACGGCGGCGGC
AAGATCGCGTTCTGTGTGCCACTCGGTGGACCAACTTTACCTTAAGCGCGCTGTCTGCTCTGCC
CCCAGAATTCCTAGCCTTTCTCCAAATTTTATCTCAAAGTCGAGAAATGGCCGATGAATTTACC
GACAATTTCAACTATCCTGAACGACAGTGGGACCGCTTCTCTTCCCCTGAACGCATTGGACAAT
GGTGCAGTCAGTTTGCACCCACCATTGCTGCCTAATGGCAATAGTGGAGGTAATCGAATGACCT
TAAAAAAAGATATGGCTGTCGACATTGATAGTACCAATTTTCGTCAAGCAGTGGCTTTATTCGC
TACGGGTATCGCTGTTCTAAGCGCCGAGACAGAAGAGGGCGATGTGCATGGCATGACTGTTAAT
AGCTTCACAAGCATTAGTCTCGACCCACCTACTGTTATGGTGAGTTTGAAATCCGGCCGAATGC
ATGAATTGTTGACTCAAGGAGGGCGTTTCGGTGTGTCTTTGCTCGGTGAAAGCCAGAAAGTGTT
TAGTGCCTTCTTCTCCAAGCGGGCAATGGACGACACACCACCCCCCGCCTTCACTATCCAAGCC
GGATTGCCAACGTTACAGGGAGCCATGGCTTGGTTTGAATGTGAAGTCGAAAGCACTGTGCAAG
TTCATGATCACACCTTGTTTATTGCCCGTGTCAGCGCTTGTGGCACCCCGGAAGCCAACACACC
TCAACCCTTGCTCTTCTTTGCCTCCCGCTATCATGGTAATCCCTTGCCCCTGAACTAAGGATCC